ASC-TR-97-5004

AN EVALUATION OF THE C-17A FORWARD LOADMASTER STATION INTERFACE

Martin N. Anesgart, Ph.D.

MISSION SYSTEMS IPT
C-17 SYSTEMS PROGRAM OFFICE
ASC/YC (AV/MS)

JANUARY 1997

FINAL REPORT FOR 1 JUNE 1996 - 31 DECEMBER 1996

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

DCS FOR INTEGRATED ENGINEERING AND TECHNICAL MANAGEMENT
AERONAUTICAL SYSTEMS CENTER
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-6503

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied said drawings, specifications or other data, is not regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

This report is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Lawrence E. Fielding

LAWRENCE E. FIELDING
Mission Systems IPT Co-Lead
C-17 System Program Office

Roger A Panton

ROGER A. PANTON
Chief of Engineering
C-17 System Program Office

If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization, please notify ASC/YC (AV/MS), Bldg 558, 2590 Loop Road West, Wright-Patterson AFB, OH 45433-7105 to help maintain a current mailing list.

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January 1997	3. REPORT TYPE AND DATES COVERED FINAL 1 June - 31 December 1996

4. TITLE AND SUBTITLE

An Evaluation of the C-17A Forward Loadmaster Station Interface

5. FUNDING NUMBERS

6. AUTHOR(S)

Martin N. Anesgart, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Missions Systems IPT
C-17 System Program Office
Aeronautical Systems Center
Wright-Patterson AFB, OH 45433-7105

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

C-17 System Program Office
Aeronautical Systems Center
Air Force Materiel Command
Wright-Patterson AFB, OH 45433-7105
ASC/YC (AV/MS), Attn: Martin Anesgart (937) 255-5427

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ASC-TR-97-5004

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION IS UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

C-17 Human Factors conducted an evaluation of the C-17A Forward Loadmaster Station (FLS) Interface in the Loadmaster Simulator at Altus AFB, OK from 29 July to 1 August 1996. By executing five training scenarios containing malfunctions, ten airdrop certified loadmasters provided information on the quality of the Station interface through rating panels, annunciators and switches on six point scales of adequacy and through responding to questions involving "Situation Awareness" (SA). Although the loadmasters generally viewed their interaction with the FLS as adequate, a repeated measures analysis on a composite of ratings and SA supported the contention that as the number and complexity of task elements increased, the ability of the FLS to aid the loadmaster in performance of the mission decreased. A follow-on ranking procedure of the average ratings demonstrated that the FLS did not give clear and full information for all phases of airdrop. Uncertainty was the primary factor contributing to a potential, serious diminishment of the loadmaster's capacity to respond in anomalous situations. Suggestions for improvements included decluttering of displays and simplication of switch actuation sequences.

14. SUBJECT TERMS

15. NUMBER OF PAGES 42

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENT

The author, much appreciating their incisive comments, extends his gratitude to Messrs. Lawrence Fielding and Melvin Santiago for their careful review of this manuscript.

1. EXECUTIVE SUMMARY

Unlike other military transport aircraft design of the The C-17A cargo compartment permits operation of the cargo compartment by only one loadmaster. Operation by one loadmaster necessitated the installation of an advanced control center known as the Forward Loadmaster Station (FLS). Workload had been a natural concern, but in extensive testing on the ground and in flight test, workload proved to be within acceptable parameters. However, workload measures can not address the quality of the interface between the loadmaster and the Station. Nor do they typically address as part of the interface "Situation Awareness" (SA) to borrow a term from the pilot community. The quality of the interface figures prominently in the timely detection and apt response to malfunctions as well as in the possible diminishment of operator error.

An assessment of the FLS interface as presented in this report included measures of interface quality through the use of nine six point rating scales similar to that provided for C-17A cockpit ratings (separate from workload ratings) and a SA measurement technique developed with the Crew Station Evaluation Facility of the Crew System Branch, Engineering Directorate, WPAFB, OH. The evaluation took place at the Loadmaster System Simulator located at Altus AFB, OK with ten airdrop certified loadmasters.

The major objective of this study targeted potential weaknesses and areas for possible improvement. Accomplishing the objective entailed performance by the loadmaster participants of five training scenarios each having distinct malfunctions: parachute deployment mechanism and right lock failure; loose platform; drogue chute failure; tow release mechanism failure; and gate release mechanism failure for the Container Delivery System. After running a scenario a participant either rated the interface on the six point scales or answered SA questions.

The loadmasters rated the FLS overall as adequate. However, a composite measure based on the ratings and SA answers as analyzed through a repeated measures statistical technique provided strong support that as the attention requirements increased for the loadmaster the quality of the interaction that the loadmaster had with the FLS decreased. A ranking procedure based on the average ratings pointed to areas that could be improved such as in providing clear and full information for the loadmaster during all phases of airdrop and response to malfunctions.

2. INTRODUCTION

The C-17A cargo compartment had been developed such that one loadmaster could operate within it. Enabling one loadmaster to perform all necessary tasks required a command center, the Forward Loadmaster Station(FLS). Appendix A depicts the displays, annunciators and switches of the FLS. While a subjective workload measure had indicated that loadmasters could accomplish their tasks without mental/physical overload and undue stress, no extensive analysis was available on the quality of the interface that the FLS provided. The quality of the interface was implicated in a mishap and through some loadmaster concerns.

Describing the interface is difficult. It does not just consist of displays, switches and annunciators and the operator looking, pushing, turning or flipping. It consists of the loadmaster's model or representation of the cargo compartment that affects and is in turn affected by the loadmaster's interaction with the FLS and by mission parameters. Evaluation of the interface thus depended on not only on how the loadmasters rated the various components of the FLS, but also on their "situation awareness," enhanced or inhibited by the FLS.

A study of the C-17A Forward Loadmaster Station was conducted in the Loadmaster System (LS) simulator at Altus Air Force Base, Oklahoma, 29 July - 1 August 1996. The study effort included participants from the C17 SPO, the U.S. Air Force, and McDonnell Douglas. Ten current and airdrop qualified C-17 loadmasters participated in the study, each performing five airdrop scenarios. Data collection consisted of responses to "Situation Awareness" (SA) questions and ratings of panels and annunciators. The ratings are described in the section on Experimental Design and Statistical Analysis. "Situation Awareness" was adopted from the pilot community and is defined in a *Handbook for Conducting Pilot-In-The-Loop Simulation for Crew Station Evaluation* (Lehman and Jenkins, 1990).

3. PURPOSE

The purpose of this study was to assess the human-machine interface at the Forward Loadmaster Station and identify weaknesses. If identified, weaknesses will be further investigated and remedies proposed in a future C-17 System Program Office Producibility Enhancement/Performance Improvement (PE/PI) Program Task Plan (PTP).

4. TEST FACILITY

The Loadmaster System simulator at Altus AFB, Oklahoma is an operational replica of the C-17A forward loadmaster station. It includes a video display which provides visual animation of events that would occur in the cargo compartment, and an over-the-shoulder instructor station. Some differences---not serious enough to affect the results---exist between the simulator and the C-17A Aircraft. These are listed in Appendix B.

5. TEST DESCRIPTION and PROCEDURE

A total of ten loadmasters participated in the study. Loadmasters performed five training scenarios noted in the section below, using the airdrop procedures from C-17 Airlift Operations, AMCI 11-217, Vol. 24, Annex B (Interim), 1 June 1996.

There were a total of fifty half-hour simulator blocks or sessions. The presentation of the blocks were randomized. Loadmasters were randomly assigned to these blocks with the constraint that each loadmaster had to provide panel/annunciator ratings and answer SA questions for each of the five scenarios. The block presentation order is shown in Appendix C.

Each half-hour session consisted of two parts. The first fifteen minutes were dedicated to completing one scenario and rating the panels and annunciators or answering SA questions. During the next fifteen minutes of the session, the loadmaster completed a second scenario, and either rated the panels and annunciators or answered SA questions (whatever they had not done after the first scenario).The order of presentation of the ratings and SA parts were randomized.

Following each scenario, the rating and SA questions were presented to the loadmaster on a laptop computer located near the simulator, but far enough away so the loadmaster could not look back at the station. Copies of the screens presented on the laptop for purposes of collecting loadmaster responses are included in Appendix D, which includes the ratings scales and their anchor definitions as well as the SA items.

6. EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS

Both ratings and SA portions of the evaluation were set into the following framework. Nine subject-matter experts were asked to rate ten mission training scenarios on four dimensions: involvement, complexity, error, and recovery, using a six point scale of "adequacy" (whose anchors and definitions are presented in Appendix E). Based on their ratings, five airdrop scenarios were chosen containing malfunctions that were best distinguishable on the four dimensions. The five chosen scenarios formed five degrees or levels (*low, medium low, medium, medium high, and high*) for each of the following dimensions:

Involvement - the degree to which the loadmaster would be expected to concentrate on viewing and interpreting the panel displays.

Complexity - the degree to which the loadmaster would be expected to concentrate on actuating the sequence of switches.

Error - the degree to which the loadmaster would be expected to misinterpret/misread a display, or actuate an incorrect switch or sequence of switches.

Recovery - the degree to which the loadmaster would be expected to overcome an error or correct a malfunction.

Each dimension representing an "independent" variable and each scenario representing a level of the variable are displayed in Table 1.

TABLE 1
LEVELS OF DIMENSIONS BY SCENARIOS

DIMENSION SCENARIO	INVOLVEMENT	COMPLEXITY	ERROR	RECOVERY
PDM&RL	medium high	medium	medium low	medium low
Loose Platform	high	high	high	low
Drogue Chute	low	low	low	high
TRM	medium	medium low	medium	medium
GRM	medium low	medium high	medium high	medium high

The five malfunctions were:

1. Sequential heavy equipment airdrop with Parachute Deployment Mechanism and Right Lock failure (PDM&RL)—lesson 311.06.09, training profile 1;
2. Heavy equipment airdrop with Loose Platform—lesson 311.06.07, training profile 2, second drop;
3. Heavy equipment airdrop with Drogue Chute failure—lesson 311.06.07, training profile 2, fourth drop;

4. Sequential heavy equipment airdrop with Tow Release Mechanism (TRM) failure—lesson 311.06.09, training profile 2; and
5. Container Delivery System (CDS) airdrop with Gate Release Mechanism (GRM) failure—lesson 311-03-21, training profile 1, second drop.

The experimental design was 5 x 4 ANOVA with repeated measures. As indicated the four dimensions were the "independent" variables, with the malfunctions being the five levels of each variable.

The method of scoring for each dimension follows:

Involvement - measured by the composite or summated average of the adequacy ratings on the four scales of viewability, readability, interpretability and decision facilitation minus SA discrepancy scores (the number or discrepancies between what occurred and what the loadmaster indicated as occurring, divided by the total possible of 7);

Complexity - measured by the composite or summated average of the adequacy ratings on the six scales of viewability, reach, actuation-force, actuation-movement, feedback and stability, minus SA discrepancy scores;

Error - measured by the composite or summated average of all nine scales minus SA discrepancy scores;

Recovery - measured by the composite or summated average of all nine scales plus the complements of SA discrepancy scores minus SA discrepancy scores ((summated average + (1 - SA discrepancy score) - SA discrepancy score)).

SPSS for Windows (SPSS, Inc., 1993) was used to do the analysis.

The repeated measures design noted above was used to test the following hypothesis:

The greater the involvement, complexity and error and the lower the recovery, the lower the adjusted rating expected.

If such a case existed, there would be some concern that the station interface did not provide all that it could to the loadmaster in performance of the mission. If no differences existed between the variable levels, there would not be a concern, given that regardless of intensity (e.g., higher workload) brought to the station interface by a particular mission scenario, the loadmaster would be expected to perform the mission effectively and safely.

7. RESULTS

The results will be presented in two parts. The first part presents a general test of the hypothesis stated in the previous section. The second part provides a narrower focus delineating problem areas, assuming that statistical significance is found for the hypothesis.

Testing the hypothesis required combining the ratings and SA discrepancy scores into a composite. Composites have a long history of use in industrial/organizational psychology to measure "overall success" or "value to the organization." The use of a composite here reflects the strong notion that the way in which the loadmasters interact with FLS displays, switches and annunciators through their senses can not be separated from their knowledge of their environment. The composite or adjusted ratings will be analyzed at the scenario and dimension level. The second part containing unadjusted ratings will target results at the scale and panel or annunciator level. Adjusting for SA discrepancies does not make sense at this level.

7.1 *Adjusted ratings*

The composites were ranked according to mean values (a rank of 1 being the lowest value) as displayed in Table 2. Note that within each dimension, Loose Platform had the lowest mean and therefore ranked last; PDM&RL the next lowest.

TABLE 2
MEANS AND STANDARD DEVIATIONS FOR SCENARIO BY DIMENSION

MEANS/STDs				
	DIMENSION			
SCENARIO	Involvement	Complexity	Error	Recovery
PDM&RL	4.557/.737	4.596/.660	4.620/.634	5.377/.693
Loose platform	4.095/.740	4.567/.764	4.587/.716	5.315/.843
Drogue chute	4.593/.639	4.676/.738	4.869/.741	5.712/.765
TRM	4.659/.668	4.646/.720	4.660/.691	5.532/.719
GRM	4.986/.668	4.950/.393	4.992/.502	5.992/.502
RANK MEANS				
	DIMENSION			
SCENARIO	Involvement	Complexity	Error	Recovery
PDM&RL	2	6	7	17
Loose platform	1	3	4	16
Drogue chute	5	11	12	19
TRM	9	8	10	18
GRM	14	13	15	20

The ranks shown in Table 2 are translated to Table 3 below comparing them to those predicted from ratings of subject matter expert on the five scenarios by four dimensions as shown in Table 1. The observed levels for Loose Platform fit the predicted levels exactly demonstrating that for high involvement, high complexity and high potential for error with low potential for recovery, loadmasters had the worst interaction with the Station compared to the other scenarios. Although not matching exactly the predicted levels, both PDM&RL and TRM provided some more support for the hypothesis expressed earlier.

TABLE 3
COMPARISION OF PREDICTED AND OBSERVED LEVELS*

DIMENSION SCENARIO	INVOLVEMENT	COMPLEXITY	ERROR	RECOVERY
PDM&RL	*medium high* medium high	*medium* medium high	*medium low* medium high	*medium low* medium low
Loose Platform	*high* high	*high* high	*high* high	*low* low
Drogue Chute	*low* medium	*low* medium low	*low* medium low	*high* medium high
TRM	*medium* medium low	*medium low* medium	*medium* medium	*medium* medium
GRM	*medium low* low	*medium high* low	*medium high* low	*medium high* high

*the levels in italics are from Table 1, the predicted variable levels.

Statistically the hypothesis was further supported by the repeated measures. The main effects for scenario and dimension were both significant at alpha=.05 (as based on three multivariate tests, Hotellings, Wilks and Roys, provided by SPSS). For the relatively problematic Loose Platform, as demonstrated in Table 4, its average composite score was significantly worse than the other scenarios for all four dimensions. Complexity was significant for PDM&RL and had a low rank as shown in Table 2 above.

TABLE 4
REPEATED MEASURES FOR SCENARIO BY DIMENSION*

	DIMENSION			
SCENARIO	Involvement	Complexity	Error	Recovery
PDM&RL		11.9948, $p<.05$		
Loose platform	10.6474, $p<.05$	254.490, $p<.05$	34.2721, $p<.05$	19.3517, $p<.05$
Drogue chute		11.2895, $p<.05$		
TRM			30.2084, $p<.05$	7.0029, $p<.05$
GRM			15.9421, $p<.05$	

*transformed variables were used to target differences between each cells and average of all the other cells as well as differences across each scenario and differences down each dimension.

7.2 *Unadjusted Ratings*

The analysis for the adjusted ratings indicated that as the number or complexity of task elements in a cargo situation increases, the quality of the interaction that the loadmaster has with the FLS decreases. To narrow in on what part of that interaction is affected, this section will go beyond the scenario and dimension levels to the scale and panel/annunciator levels. Table 5 displays the average (or mean) ratings made by each loadmaster as well each of their SA discrepancy score (SA-D). Note that two of the loadmasters had relatively high SA-D, but loadmaster 5 had ratings of the FLS that were among the lowest, while loadmaster 8 had ratings among the highest. In fact, there was no significant rank order correlation (p>.05) between SA-D and the mean ratings. And neither were there significant correlations between C-17 and overall experience with the average ratings. It should be noted here that there is probably some restriction of range on at least one of the experience variables. All the loadmasters were male and generally very experienced. A restricted range on at least one of the correlated variables can result in a lower correlation than would be likely in the population at large. It may be that with a much larger and diverse sample experience would be a factor in the ratings (as well as SA-D).

TABLE 5
OVERALL RATINGS NOT ADJUSTED FOR SA DISCREPANCY SCORES

	C-17 Exp.	Overall Exp.	Mean	Std.	Coef. Var.+	N	SA - D*
Loadmaster1	1.5	8	5.3125	0.9326	21.63%	96	0.2500
Loadmaster2	3	16	4.0729	0.8491	27.63%	96	0.2857
Loadmaster3	5	15	5.4479	0.7380	16.59%	96	0.0715
Loadmaster4	2.5	5	4.7292	0.5126	13.75%	96	0.1429
Loadmaster5	1	8.5	4.1354	0.4500	14.35%	96	0.3929
Loadmaster6	1	11	4.6042	0.7466	16.22%	96	0.1072
Loadmaster7	7	22	4.0938	1.2318	39.82%	96	0.0357
Loadmaster8	6	17.5	5.2917	0.4569	10.65%	96	0.3215
Loadmaster9	3	19	4.9479	0.4441	11.25%	96	0.1786
Loadmaster10	4	18	4.7917	0.4794	12.64%	96	0.2143
Spearman Rank Order Correlation	with Mean = -0.3030ns	with Mean = 0.0788ns					with Mean =-0.2364ns
Entire Population			4.8427	0.9118	23.73%	960	*out of 4 scen.

+this column is the coefficient of variation in percent: the standard deviation (Std.) divided by the mean, the result multiplied by 100.

With the unadjusted ratings all values are in the adequate range of the six point scale defined in Appendix D. However, the purpose of the study was to look at relative weaknesses if any exist. The analysis for the adjusted ratings indicated support for weaknesses existing. To better define those weaknesses a mean below 4.5 will be considered marginal (the six point adequacy scale has historically been used as an equal interval scale, thus scale points 1 to 6 can be thought of as intervals, e.g., mildly adequate goes from 3.5 to 4.5; very adequate from 4.5 to 5.5).

Table 6 below displays marginal values for combinations of scenario, scale and category of panel/annunciators. The information type scales, Feedback, Readability, Interpretability and Decision Facilitation made up the bulk of the marginal ratings; and the ADS lock annunciator had almost half of the categories rated marginally.

TABLE 6
UNADJUSTED RATINGS FOR SCENARIO BY SCALE BY CATEGORY

SCENARI0	SCALE	CATEGORY N=10	MEAN	STD	MIN-MAX RATING	SCALE MEAN	SCALE STD	N
PDM&RL	Viewability	GangLK Backup Panel	4.3	1.0593	3.0-6.0	4.5333	1.0080	30
PDM&RL	Reach	GangLK Backup Panel	4.4	1.2649	2.0-6.0	4.7667	1.1943	30
PDM&RL	Decision Facilitation	ADS Lock Annunciator	4.3	1.0593	3.0-6.0	4.8500	0.9881	20
Loose Platform	Viewability	GangLK Backup Panel	4.4	1.0750	3.0-6.0	4.7000	0.8739	30
Loose Platform	Readability	ADS Lock Annunciator	4.4	0.8433	3.0-6.0	4.8500	0.8751	20
Loose Platform	Interpretability	ADS Lock Annunciator	4.4	0.8433	3.0-6.0	4.8000	0.9515	20
Drogue Chute	Feedback	FWD Control Panel	4.3	0.8233	3.0-5.0			
Drogue Chute	Feedback	ADS Backup Panel	4.4	0.9661	2.0-5.0	4.3500	0.8751	20
Drogue Chute	Readability	ADS Lock Annunciator	4.4	1.0750	2.0-6.0	4.7500	0.9105	20
Drogue Chute	Decision Facilitation	ADS Lock Annunciator	4.4	0.8433	3.0-6.0	4.7500	0.7864	20
TRM	Reach	ADS Backup Panel	4.3	1.4181	1.0-6.0	4.6000	1.2312	20
GRM	Feedback	FWD Control Panel	4.4	0.6992	3.0-5.0	same	same	10

Ranking the average (mean) ratings (all loadmasters included) for 96 scenario, scale and category combinations produced average (median) ranks as shown in Table 7 below. Of the scales, Feedback had the lowest overall ranking (16.5). Viewability was the next lowest (31.25). Of the panels/annunciators, The ADS Lock Annunciator had the lowest ranking (16.5) across scales. The ADS Backup Panel and the CDS Armed Annunciator had the next lowest (both 26.5). Of the combinations of scale and category, Viewability and GANGLK Backup had the lowest ranking (5.75); Feedback/ADS Backup Panel, Readability/ADS Lock Annunciator, and Decision Facilitation/ADS Lock Annnunciator had the next lowest (at 12.75). The latter two scales plus Viewability madeup the dimension of Involvement (display concentration) and Feedback was part of Complexity (switchology concentration). Both dimensions contributed to Error, all concerns for scenarios one and two as based on analysis of the ratings adjusted for SA discrepancy scores.

TABLE 7
MEDIAN RANKS FOR SCALE, CATEGORY AND SCALE BY CATEGORY

SCALE	CATEGORY	MEDIAN RANK FOR CATEGORY	MEDIAN RANK FOR SCALE
Viewability	FWD Control Panel	53	
Viewability	ADS Backup Panel	34.75	
Viewabililty	GangLK Backup Panel	5.75	31.25
Reach	FWD Control Panel	53	
Reach	ADS Backup Panel	26.5	
Reach	GangLK Backup Panel	26	43
Actuation-Force	FWD Control Panel	64	
Actuation-Force	ADS Backup Panel	43	
Actuation-Force	GangLK Backup Panel	39.5	48
Actuation-Movement	FWD Control Panel	82.5	
Actuation-Movement	ADS Backup Panel	53.5	
Actuation-Movement	GangLK Backup Panel	67.75	73.25
Feedback	FWD Control Panel	16.5	
Feedback	ADS Backup Panel	12.75	
Feedback	GangLK Backup Panel	34.75	16.5
Stability	FWD Control Panel	82.5	
Stability	ADS Backup Panel	63.25	
Stability	GangLK Backup Panel	48	73.5
Readability	ADS Lock Annunciator	12.75	
Readability	Door NLKED Annunciator	82.5	
Readability	CDS Armed Annunciator	82.5	49.5
Interpretability	ADS Lock Annunciator	26.5	
Interpretability	Door NLKED Annunciator	73.5	
Interpretability	CDS Armed Annunciator	26.5	50
Decision Facilitation	ADS Lock Annunciator	12.75	
Decision Facilitation	Door NLKED Annunciator	73.5	
Decision Facilitation	CDS Armed Annunciator	26.5	48.25
		OVERALL	
	FWD Control Panel	64	
	ADS Backup Panel	26.5	
	GangLK Backup Panel	43	
	ADS Lock Annunciator	16.5	
	Door NLKED Annunciator	73.5	
	CDS Armed Annunciator	26.5	

8. DISCUSSION

The purpose of this study was to examine the human-machine interface for the Forward Loadmaster Station and identify weaknesses. Use of training scenarios was considered the optimum method for performing the evaluation. Five scenarios simulating five malfunctions—Parachute Deployment Mechanism & Right Lock failure (PDM&RL), Loose Platform, Drogue Chute, Tow Release Mechanism (TRM) and CDS Gate Release Mechanism (GRM) failures---provided the venue for assessing the interface on nine rating scales, ratings of which were adjusted for a score on situation awareness by each of the ten loadmasters. The scores on the nine scales were distributed among four dimensions. The dimensions Involvement, Complexity, Error and Recovery addressed components of the interface, displays for Involvement, Switches for Complexity, combination of displays and switches for Error and Recovery. The difference between the latter two was in the way the discrepancy score on "Situation Awareness" (SA) was applied as described in the section on experimental design and statistics.

8.1 *Adjusted Ratings*

Recall that adjusted ratings take "Situation Awareness" into account, and that these adjusted ratings could only be analyzed at the scenario and dimension level, not at the scale or panel/annunciator level. Recall also as noted in the experimental design and statistical analysis section that the dimensions are the "independent" variables with each scenario as one level of the variable as based on subject matter expert ratings of ten scenarios, from which five were chosen. Loose Platform had the cleanest profile: high in Involvement, high in Complexity, high in Error and Low in Recovery. Statistically, as based on repeated measures all four levels were significantly different from other variable/level combinations. Further in ranking the variable/level means, Loose Platform (with the door open) ranked the lowest for each variable as would be expected for the hypothesis expressed in the section experimental design and statistical analysis: the higher degree the loadmaster had to concentrate on the displays and switches, the greater likelihood of error and lower likelihood of recovery from error, and the lower the expected adjusted rating.

In the Loose Platform scenario the loadmaster received no indication from ADS lock status panel or any other annunciators which pallet was loose (or for that matter which pallets are actually engaged in the locks; increasing loadmaster uncertainty and potential for error as well as creating difficulty in sequencing switches far apart from the gang lock to ADS backup lock). The results for the other scenarios fit to some extent the hypothesis. While PDM&RL had the next worst score in involvement, it was higher in complexity (found to be significant, $p<.05$, by the repeated measures) and error than expected. Uncertainty is a strong condition of both scenarios (SA discrepancy was highest for the two). Knowing where a pallet is located and which locks are engaged in that pallet is critical for successful reaction to a malfunction, particularly the type of malfunctions represented by Loose Platform and PDM&RL. When an uncertain failure occurs, the loadmaster may hit a various sequence of switches related to a heavy equipment drop on a trial and error basis, complicated by the loadmaster having to readjust lock settings, for example, moving between the backup panels high in the loadmaster station and the rotary

switches on the lower part of the station. Where a malfunction is more certain and the number of switches and displays with which the loadmaster must interact is to a lesser degree, dimensions were significantly ($p>.05$) a nonproblem, e.g., Drogue Chute failure had the lowest significantly rated Complexity and GRM was expected to be medium-high in error potential, but came out low. SA was 0 for the CDS scenario, but this zero value may be a byproduct of the way SA was operationalized in this study. For CDS there are no displays to look at and only two switches to sequence, CDS gate select and CDS release; yet the loadmasters in their comments wished to see indicators for each gate.

8.2 *Unadjusted Ratings*

The scale, Reach, had the most variability in unadjusted ratings, yet, there was no significant ($p>.05$) rank order correlation between loadmaster functional and extended functional reach and loadmaster ratings, r_s=.22 and .02, respectively. The next highest variability was Viewability, and although no significant ($p>.05$) correlation existed between loadmaster height (a proxy for better or worse for sitting eye height) and loadmaster ratings, a r_s of .36 in the low-moderate range would suggest a tendency for those of greater stature to more quickly locate a control and display in general or specifically, more easily read labels on switches on the backup panels at a greater height on the Forward Loadmaster Station. Viewability also relates to confusion between switches/panels that look the same.

Although ratings on all the scales were in the adequate range, ranking the scales against each other (Table 7 of the results) revealed that Feedback (or whether a control provided position or condition information) was rated lowest. Feedback was interpreted broadly by the loadmasters as their comments would suggest and included the lack of a status indication of individual platforms and locks for heavy equipment drop and of gate locations for CDS drop. This applied to the general "FWD Control Panel" (16.5) as well as more specifically to the ADS Backup Panel (12.75). For the latter, when coupled with low ranks for Readability(12.5), Interpretability (26.5) and Decision Facilitation (12.5) for the ADS Lock Annunciator suggests that the loadmaster station is lacking in providing enough indicators to the loadmaster of what is occurring in the cargo compartment.

The findings for the unadjusted ratings complement the findings with the adjusted ratings. The greater the uncertainty as represented by Feedback, Readability, Interpretability and Decision Facilitation identified as weaknesses, the more the loadmaster has to concentrate on displays and the greater the trial and error in actuating switches, cutting his or her ability to respond efficaciously in a malfunction situation where time and safety factors are critical and potential for error is high, while the capacity to overcome that error is low.

9. CONCLUSIONS AND RECOMMENDATIONS

The loadmasters generally viewed the FLS as adequate, but the analysis of their ratings and scores on SA pointed to areas that could be improved, primarily in providing clear and full information; one example being the location of a loose pallet. Uncertainty is a factor contributing to a possible lessening of the loadmaster's capacity to respond to a malfunction. The more complex the situation the greater the burden on the loadmaster and the less likely the Station will provide the needed information. The specific factors creating this uncertainty would have to be identified in follow up studies given the constraints of the present study. Yet, the data presented here as well as the comments from the loadmasters suggest that a number of improvements could be made to the Station. These include providing: a clear status of switches and could include cutting the lag time between activation of a switch and its outcome; exact pallet location and which locks are engaged in which pallet; appropriate and timely indication of a malfunction and identifying what specifically that malfunction is; and increasing an effective and timely response to a malfunction that could include decluttering displays of extraneous information for a particular mission and simplifying switch actuation sequences.

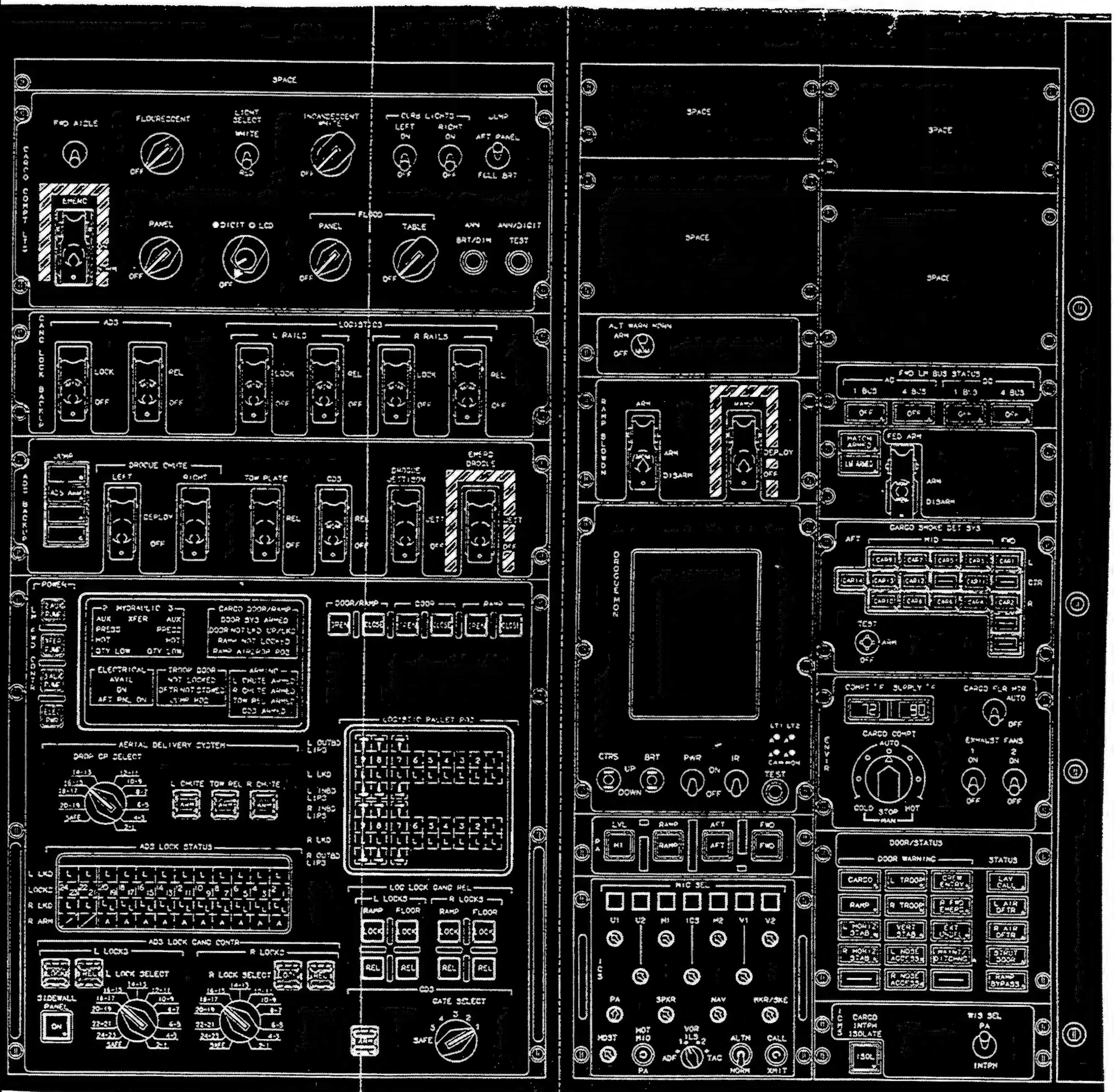
10. REFERENCES

Lehman, E.F. and Jenkins, M.(1990). Handbook for Conducting Pilot-In-The-Loop Simulation for Crew Station Evaluation (HSD-TR-90-007). Brooks AFB, Texas.

APPENDICES

APPENDIX A

The Forward Loadmaster Station Displays, Annunciators and Switches



APPENDIX B

Differences between the C-17 Loadmaster System Simulator and the C-17 Aircraft

1. For the backup gang release and lock switches, the lag time between actuation of the switches and indication that the locks have been locked or released is approximately 5 seconds in the simulator and approximately 2.5 seconds in the aircraft.
2. A video monitor with low fidelity graphics is provided in the simulator for a view of the cargo compartment.
3. The drogue monitor in the simulator does not provide an accurate display of drogue chute(s).
4. The #3 hydraulic pressure indication is present in the aircraft but not in the simulator.
5. The maintenance panel (containing OBIGGS switch, for example) is two-dimensional in the simulator.
6. The emergency drogue switch is momentary in the simulator, but positional in aircraft.
7. Door and ramp open and close push-button switches must be held in the simulator; in the aircraft, one touch begins actuation.
8. The ramp and door open and close simultaneously in the aircraft, but not in the simulator
9. The ramp and door take longer to open and close in simulator than on the aircraft.
10. The cargo compartment panel (containing static line retrieval switch, for example) is two-dimensional (on a monitor) in the simulator.
11. The hashmarks surrounding the smoke detector switch/indicator are present in the aircraft, but not in the simulator.

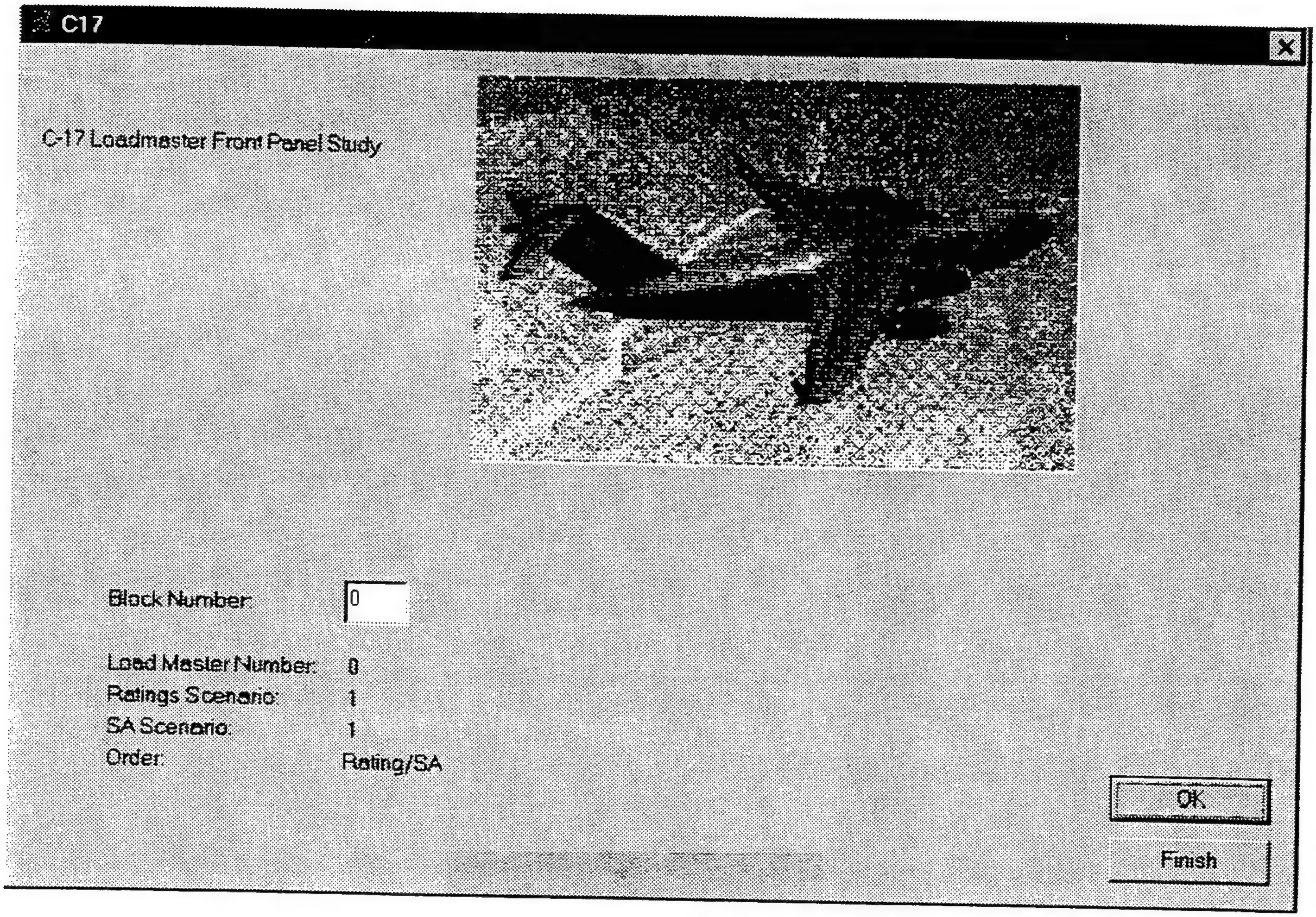
APPENDIX C

RANDOMIZED BLOCKS FOR PRESENTATION OF SCENARIOS

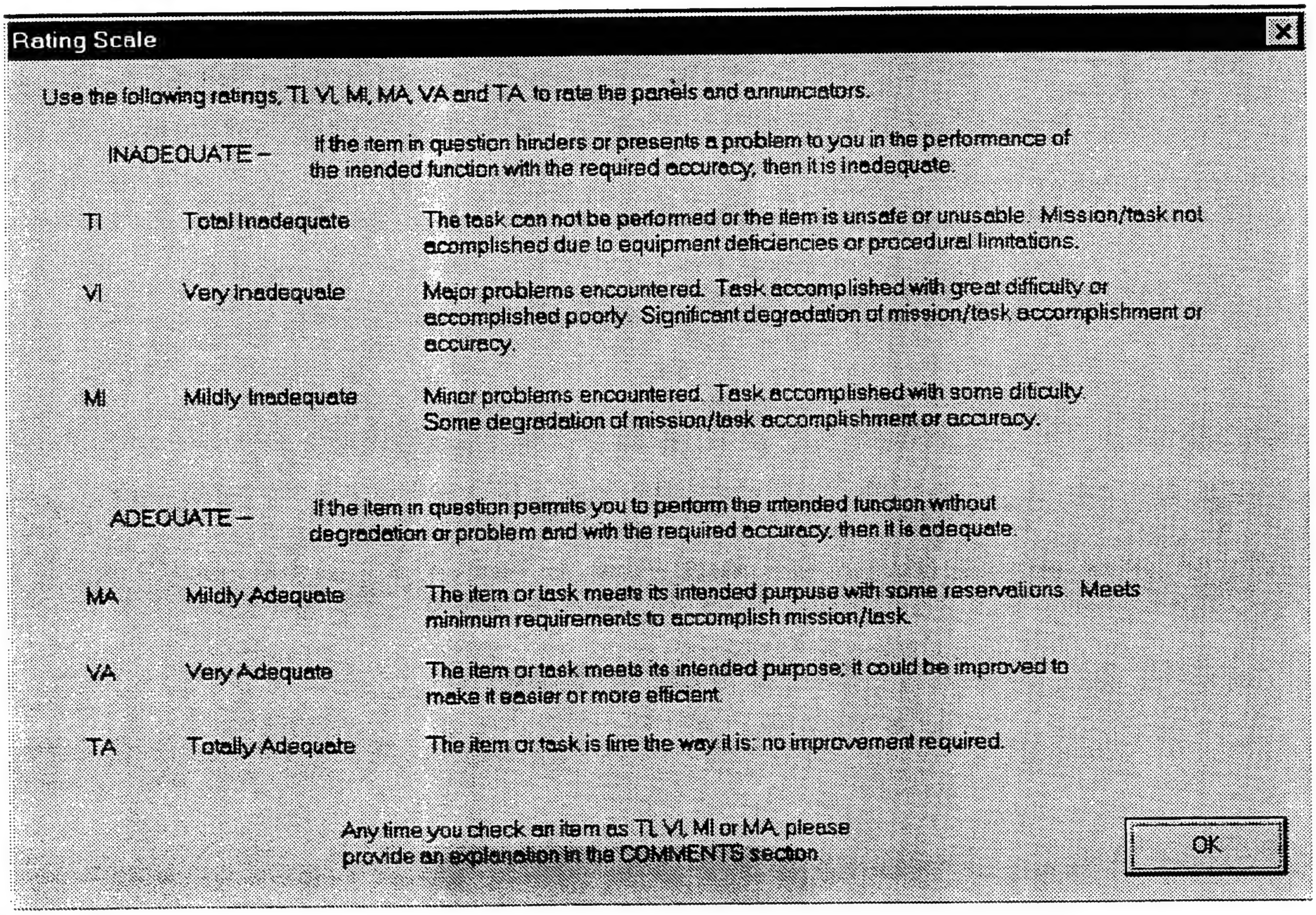
Block	Ratings	SA assess	LM #	Block	SA assess	Ratings	LM #
1	1	1	9				
				2	2	5	3
				3	2	1	5
4	1	2	10				
5	3	3	8				
6	5	1	5				
7	5	5	7				
				8	1	1	1
				9	3	1	2
				10	4	1	4
11	5	3	6				
				12	5	2	2
13	4	4	5				
				14	5	5	4
				15	3	2	9
16	4	5	9				
17	2	3	7				
18	3	1	7				
				19	4	3	10
				20	3	5	10
				21	2	3	4
				22	1	3	2
				23	5	3	3
				24	4	5	1
25	1	3	3				
26	2	5	5				
27	3	5	1				
				28	3	4	1
29	1	5	6				
30	1	4	7				
31	2	1	10				
32	5	4	2				
33	3	4	9				
34	3	2	6				
				35	1	5	8
				36	4	4	8
				37	4	2	6
				38	5	4	10
				39	2	2	8
40	5	2	9				
				41	1	4	6
				42	5	1	8
43	4	3	4				
44	2	2	1				
45	2	4	3				
46	4	1	3				
				47	2	4	7
				48	3	3	5
49	4	2	2				
				50	1	2	4

APPENDIX D

EVAUATION SCREENS PRESENTED TO THE LOADMASTERS



APPENDIX D (CONT'D)



Rating Scale

Use the following ratings, TI, VI, MI, MA, VA and TA to rate the panels and annunciators.

INADEQUATE – If the item in question hinders or presents a problem to you in the performance of the inended function with the required accuracy, then it is inadequate.

TI	Total Inadequate	The task can not be performed or the item is unsafe or unusable. Mission/task not acomplished due to equipment deficiencies or procedural limitations.
VI	Very Inadequate	Major problems encountered. Task accomplished with great difficulty or accomplished poorly. Significant degradation of mission/task accomplishment or accuracy.
MI	Mildly Inadequate	Minor problems encountered. Task accomplished with some dificulty. Some degradation of mission/task accomplishment or accuracy.

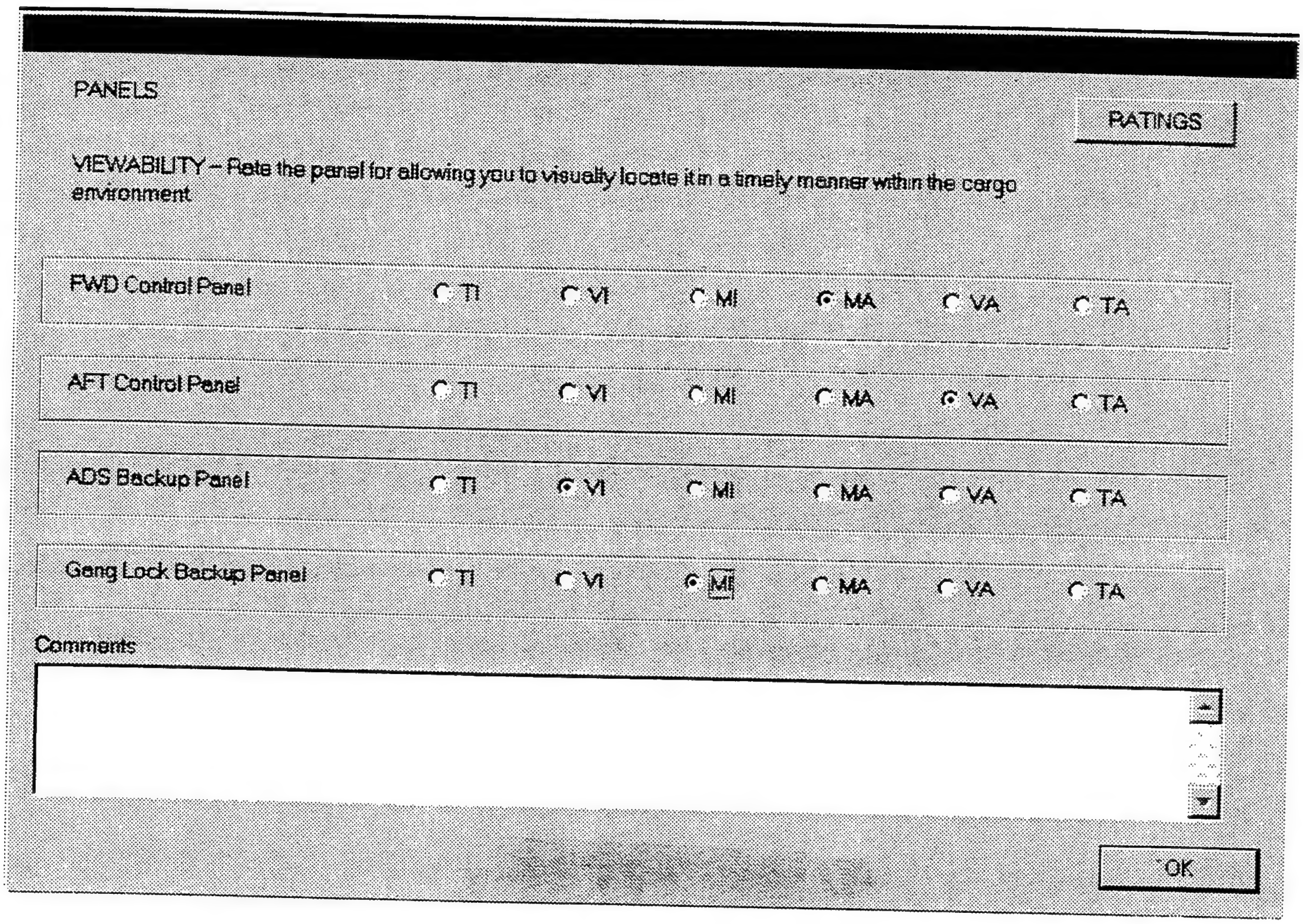
ADEQUATE – If the item in question permits you to perform the intended function without degradation or problem and with the required accuracy, then it is adequate.

MA	Mildly Adequate	The item or task meets its intended purpose with some reservations. Meets minimum requirements to accomplish mission/task.
VA	Very Adequate	The item or task meets its intended purpose; it could be improved to make it easier or more efficient.
TA	Totally Adequate	The item or task is fine the way it is; no improvement required.

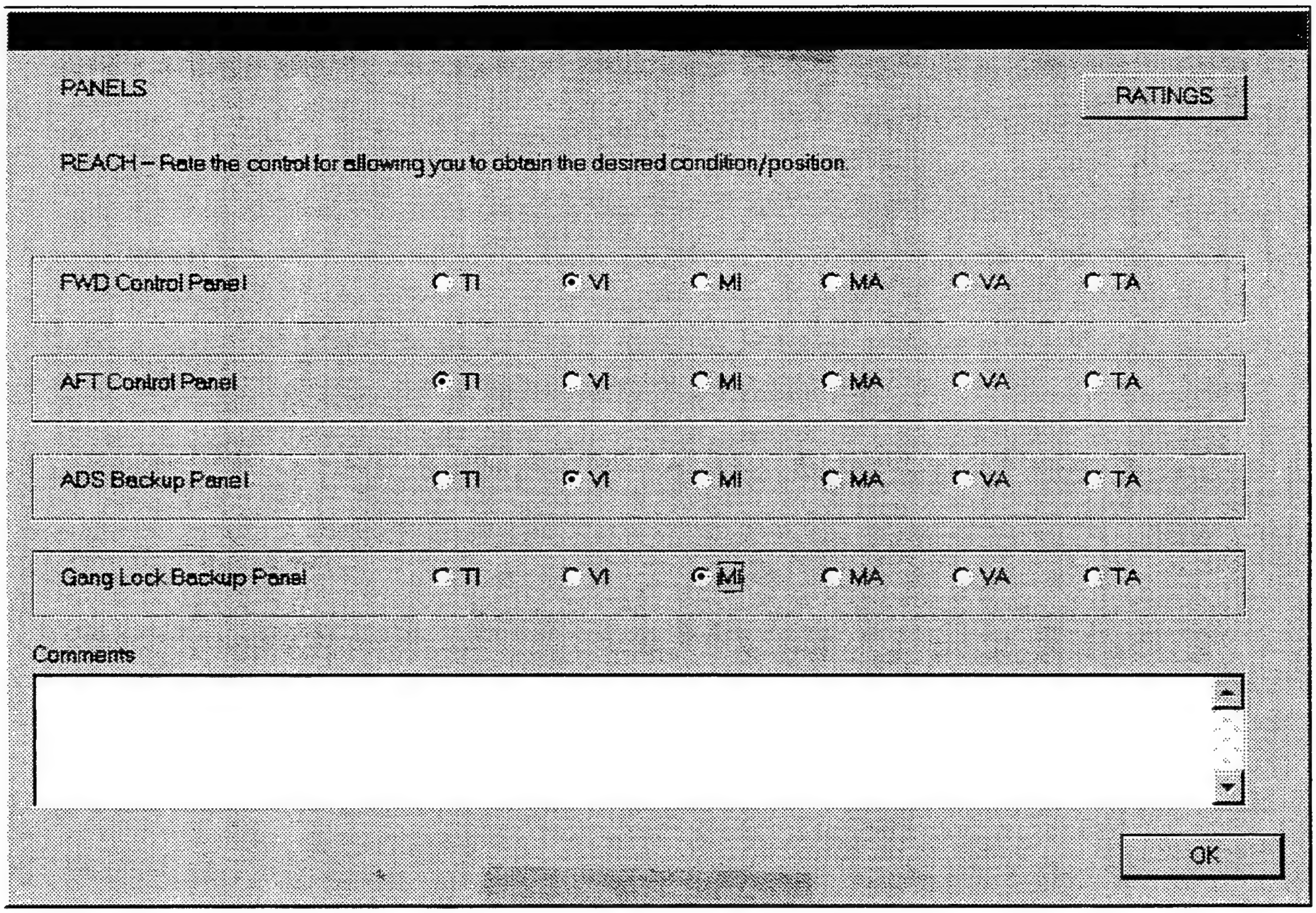
Any time you check an item as TI, VI, MI or MA, please provide an explanation in the COMMENTS section.

OK

APPENDIX D (CONT'D)



APPENDIX D (CONT'D)


PANELS

RATINGS

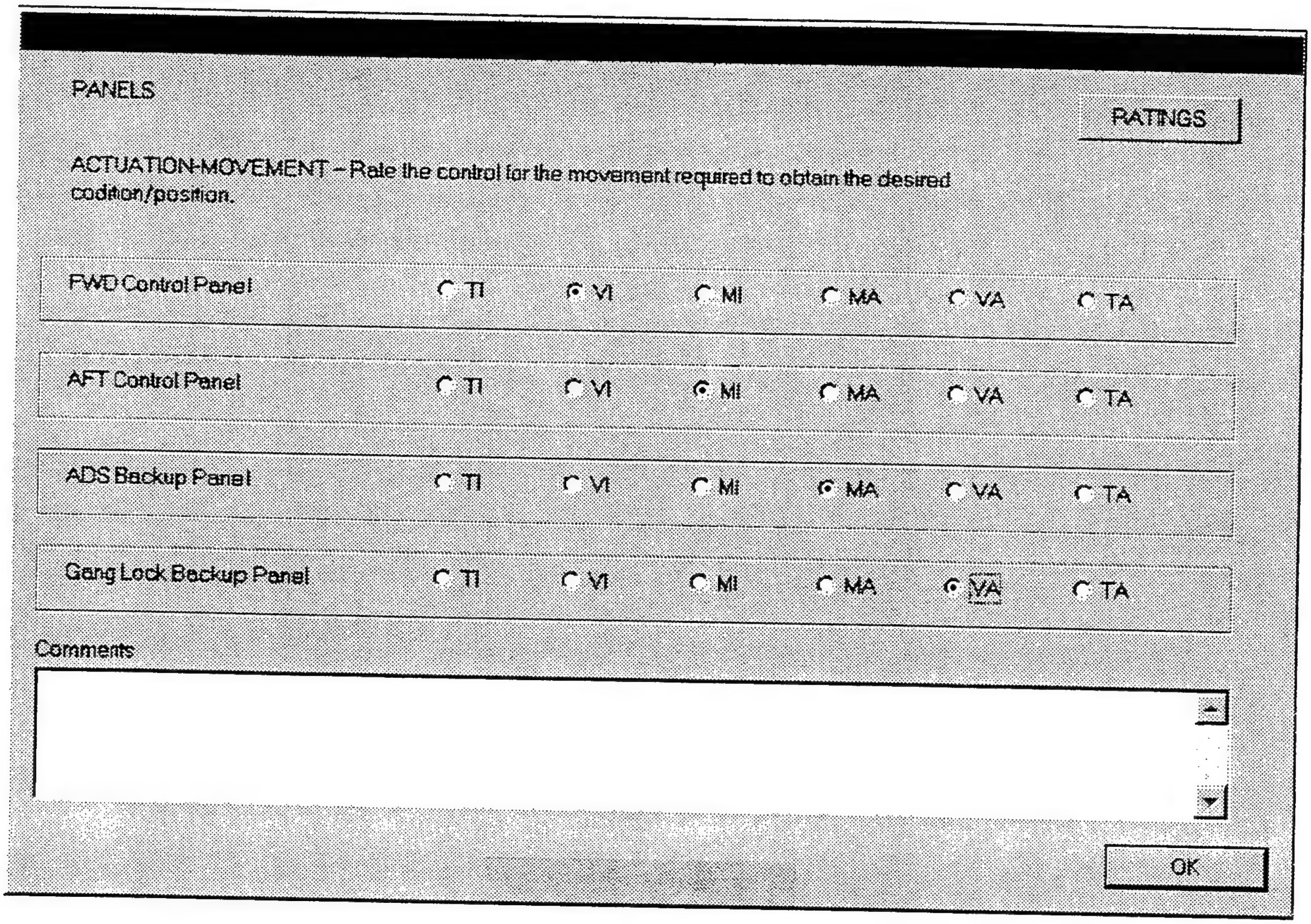
REACH – Rate the control for allowing you to obtain the desired condition/position.

FWD Control Panel	TI	VI	MI	MA	VA	TA
AFT Control Panel	TI	VI	MI	MA	VA	TA
ADS Backup Panel	TI	VI	MI	MA	VA	TA
Gang Lock Backup Panel	TI	VI	MI	MA	VA	TA

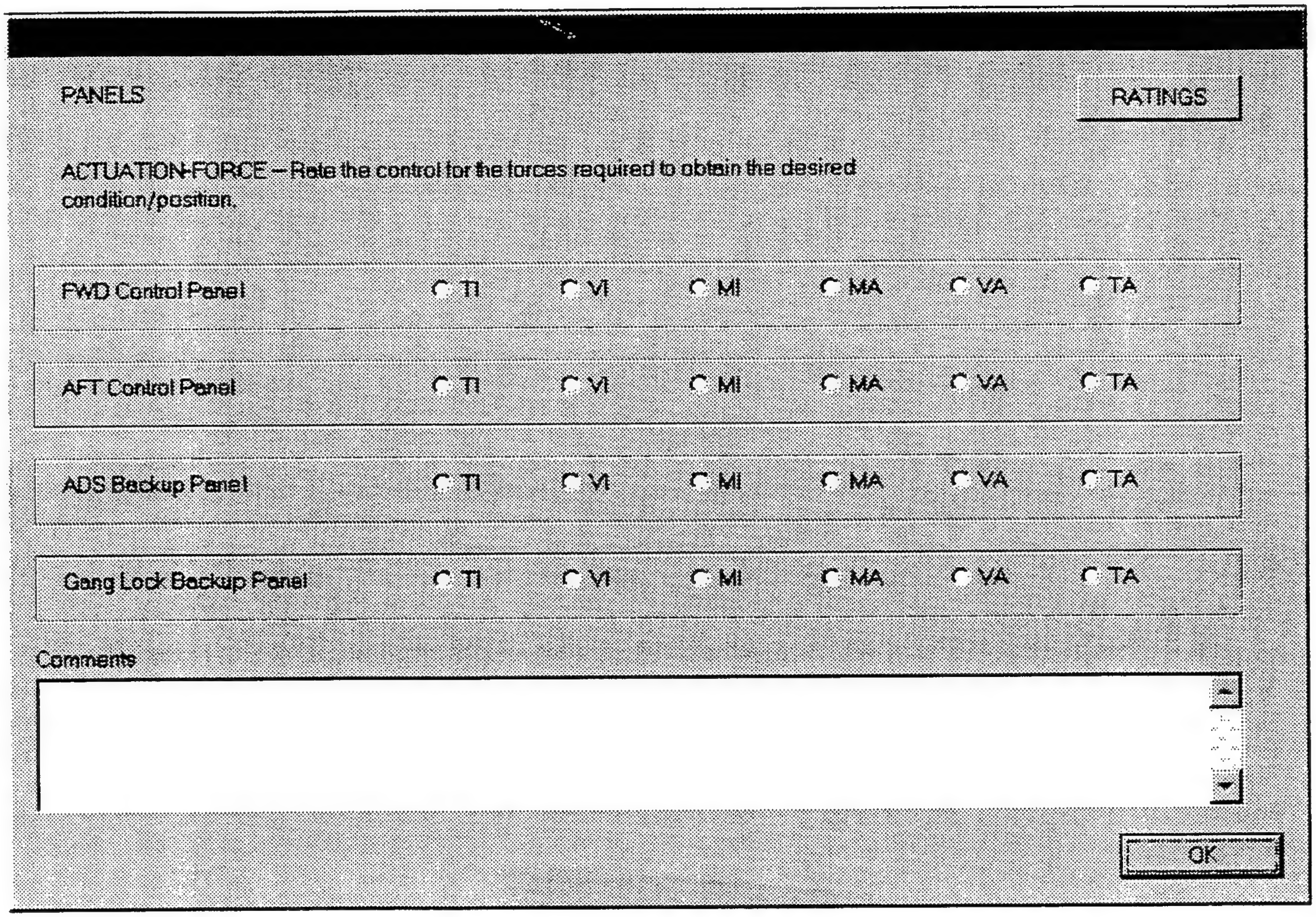
Comments

OK

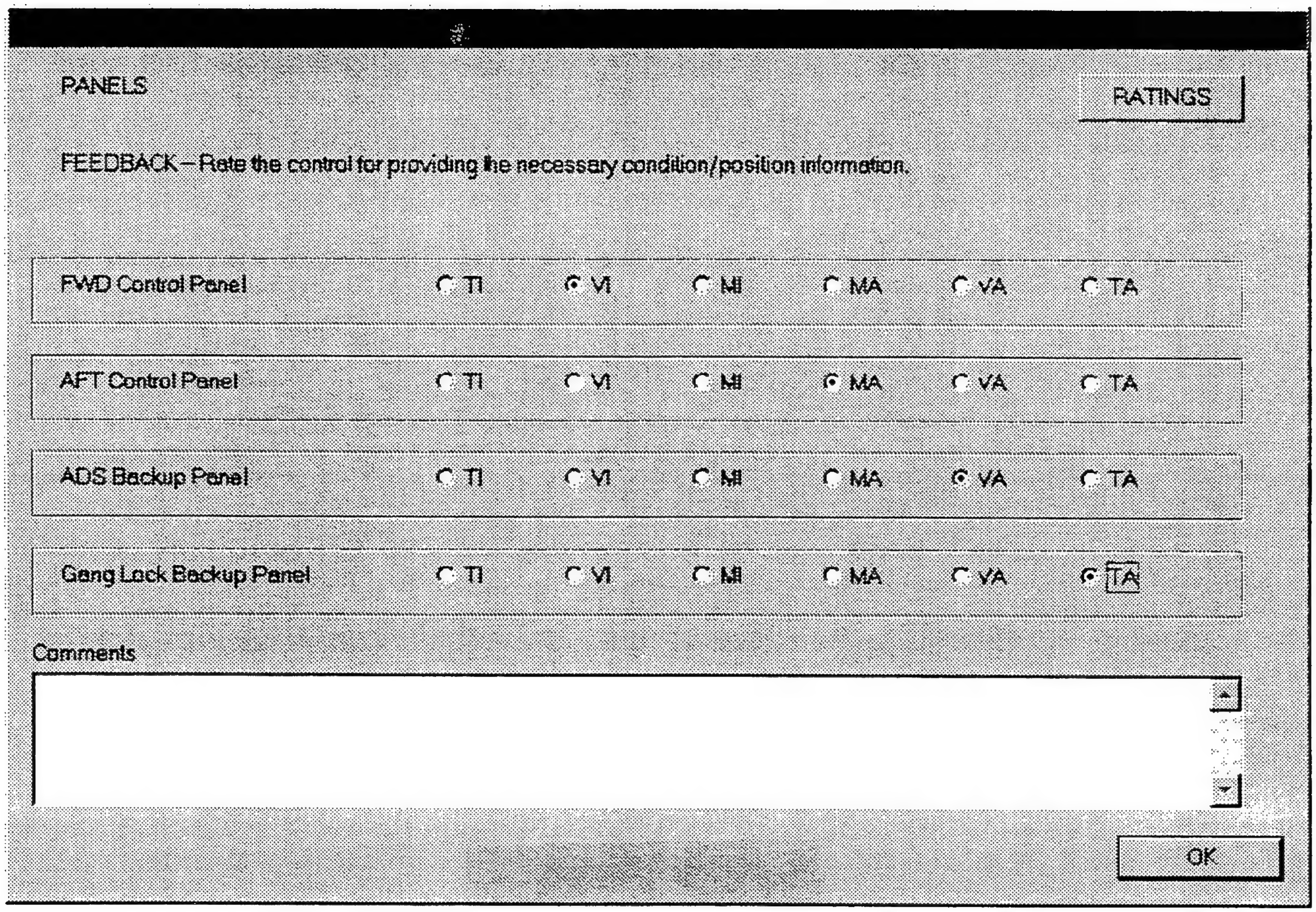
APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



PANELS

RATINGS

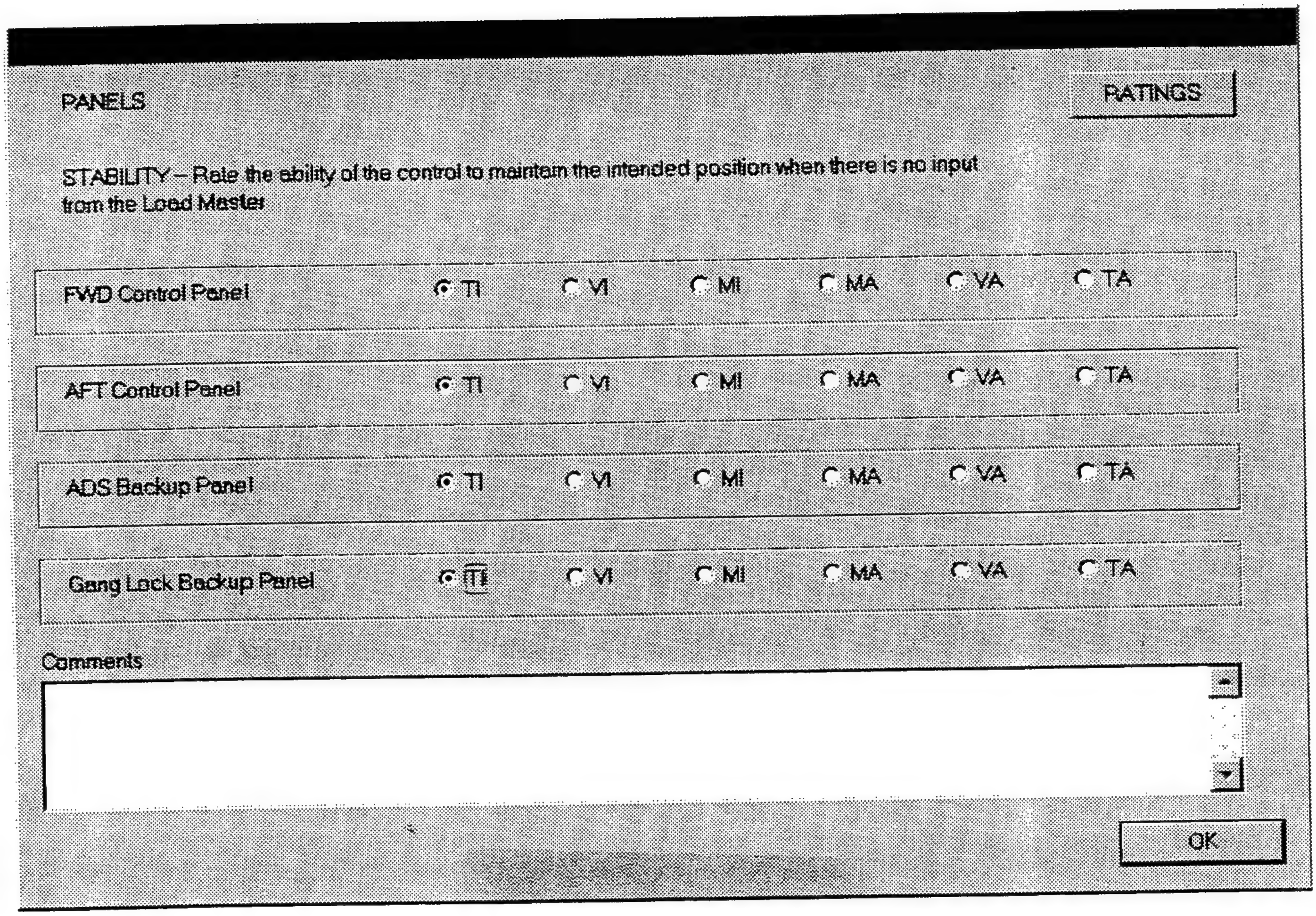
FEEDBACK – Rate the control for providing the necessary condition/position information.

FWD Control Panel	TI	VI	MI	MA	VA	TA
AFT Control Panel	TI	VI	MI	MA	VA	TA
ADS Backup Panel	TI	VI	MI	MA	VA	TA
Gang Lock Backup Panel	TI	VI	MI	MA	VA	TA

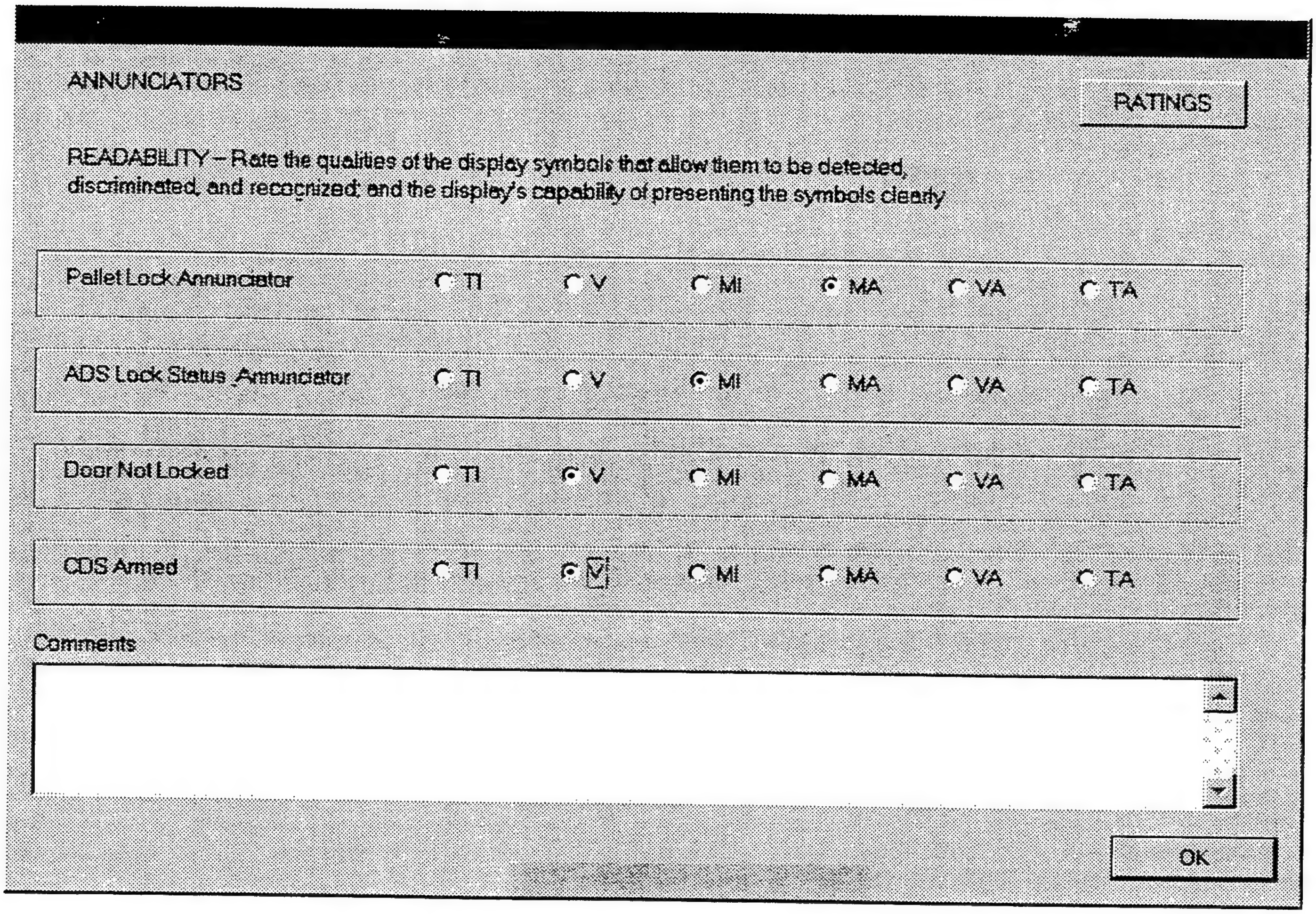
Comments

OK

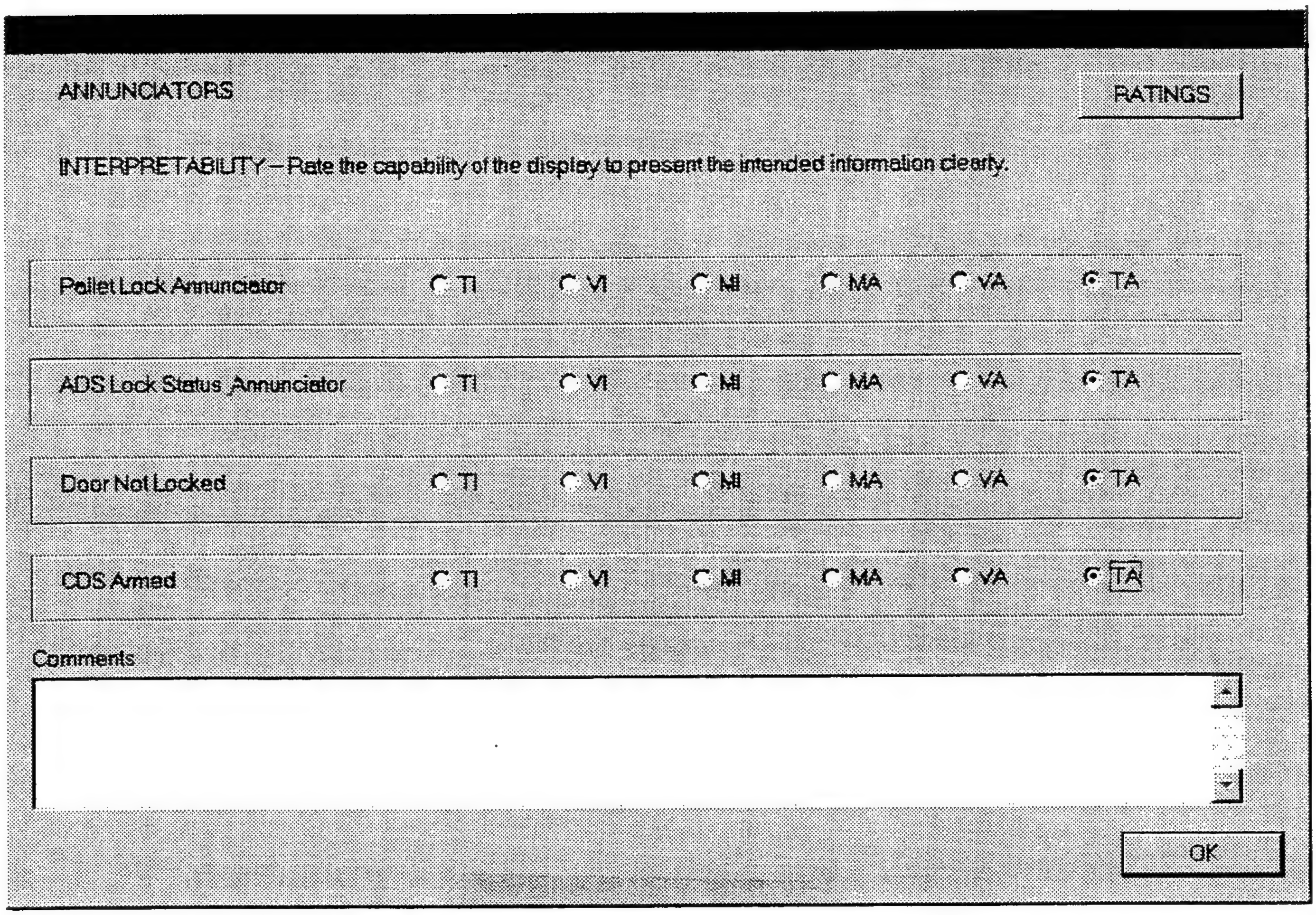
APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



ANNUNCIATORS

RATINGS

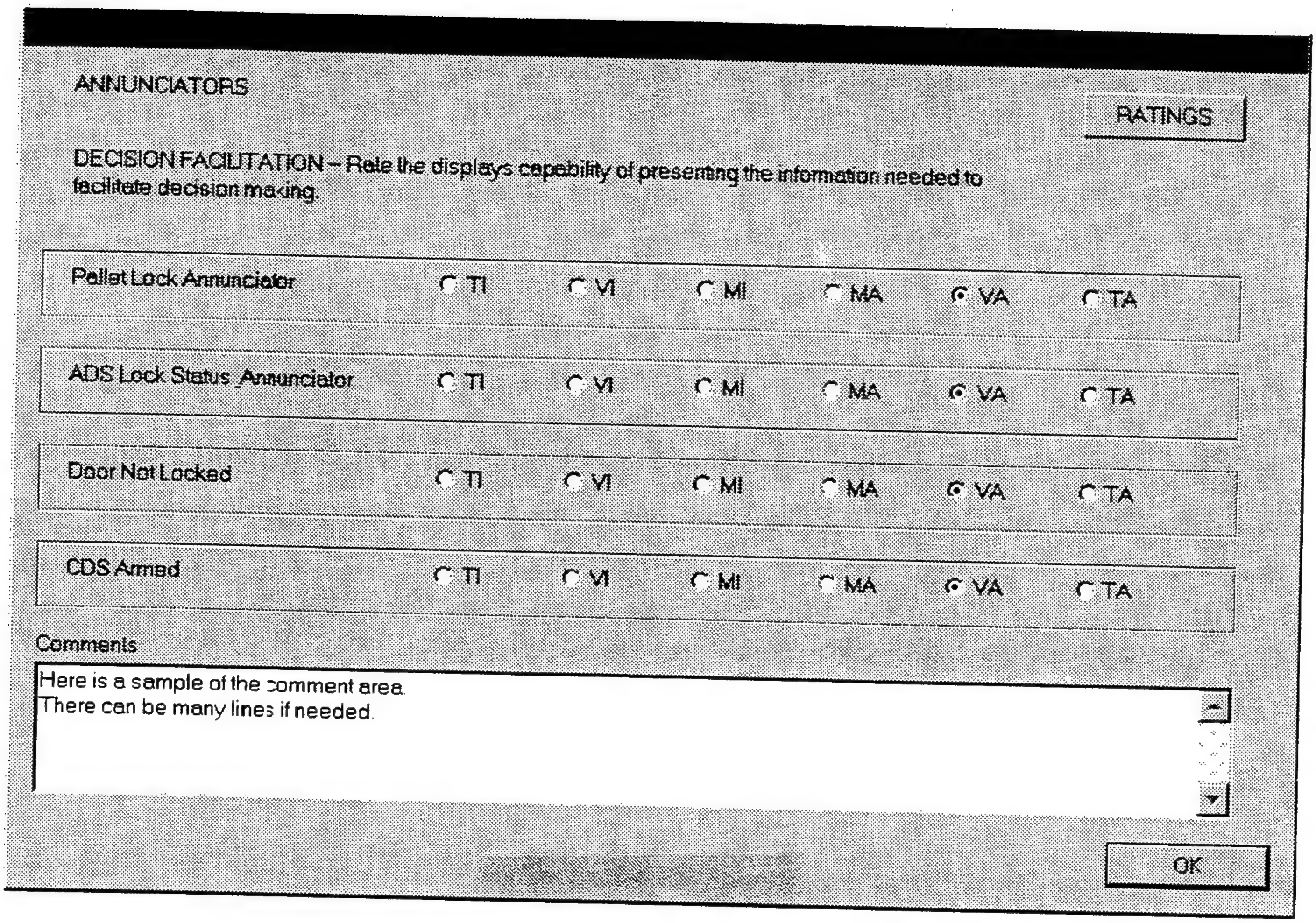
INTERPRETABILITY – Rate the capability of the display to present the intended information clearly.

Pallet Lock Annunciator	TI	VI	MI	MA	VA	TA
ADS Lock Status Annunciator	TI	VI	MI	MA	VA	TA
Door Not Locked	TI	VI	MI	MA	VA	TA
CDS Armed	TI	VI	MI	MA	VA	TA

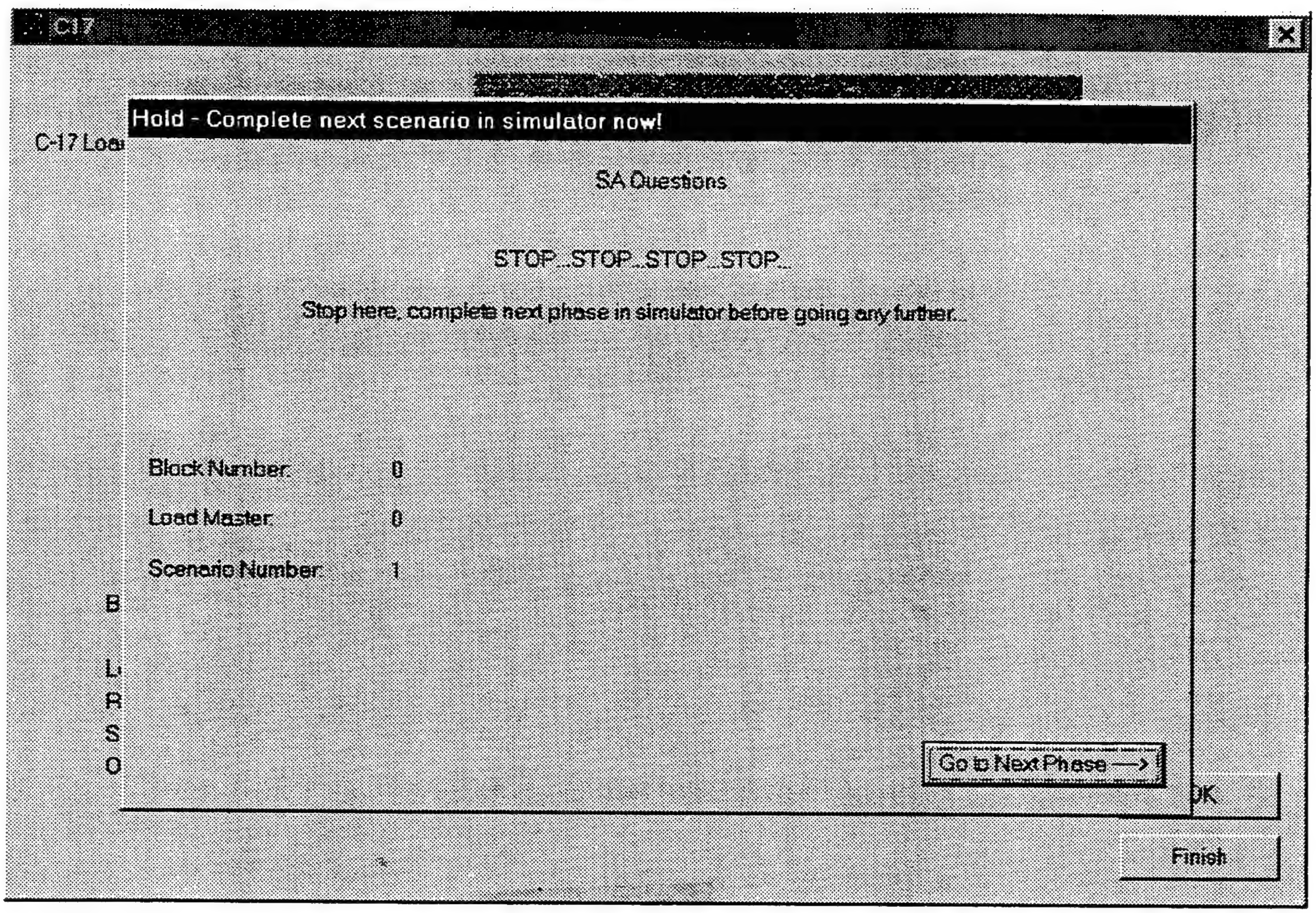
Comments

OK

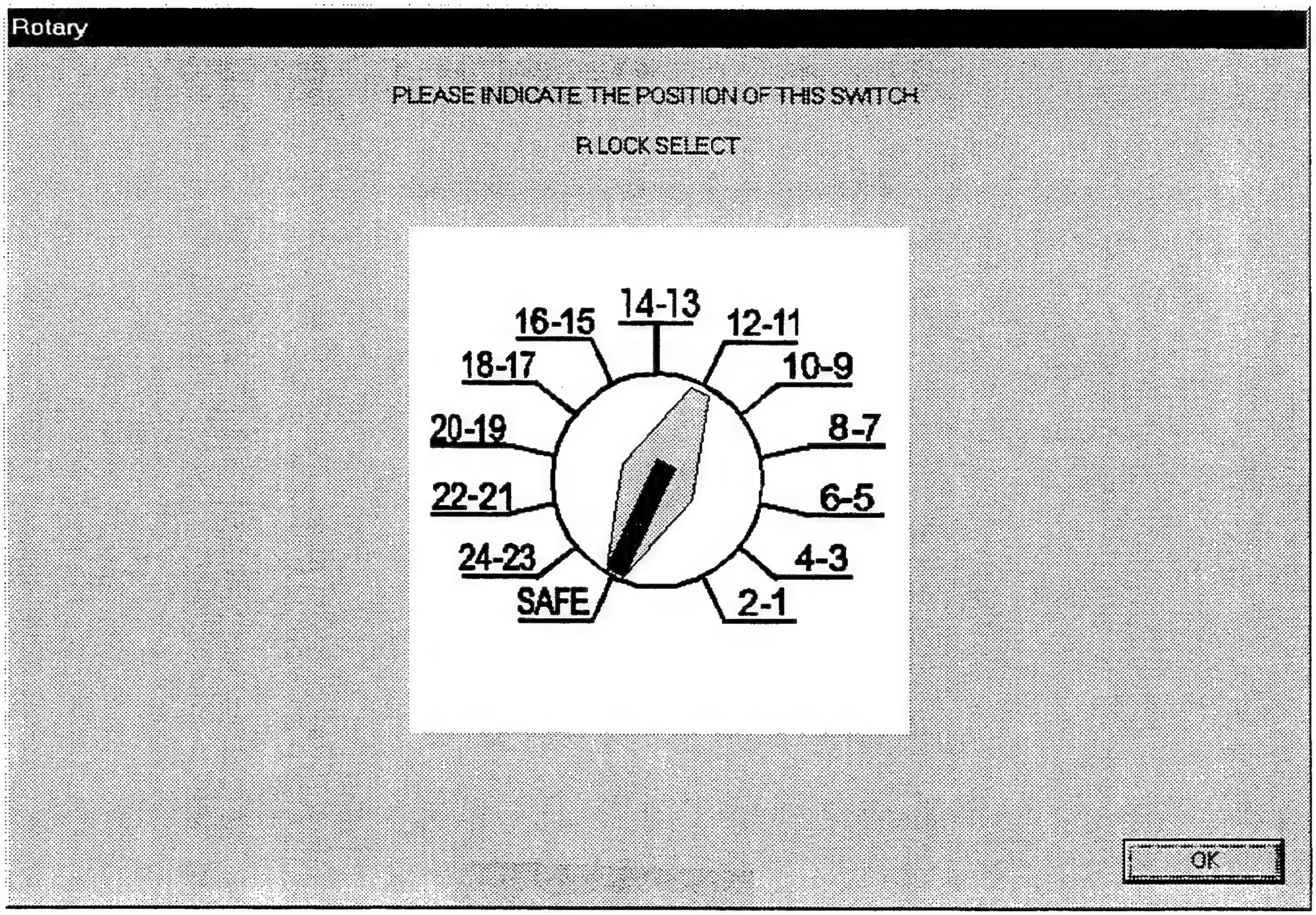
APPENDIX D (CONT'D)



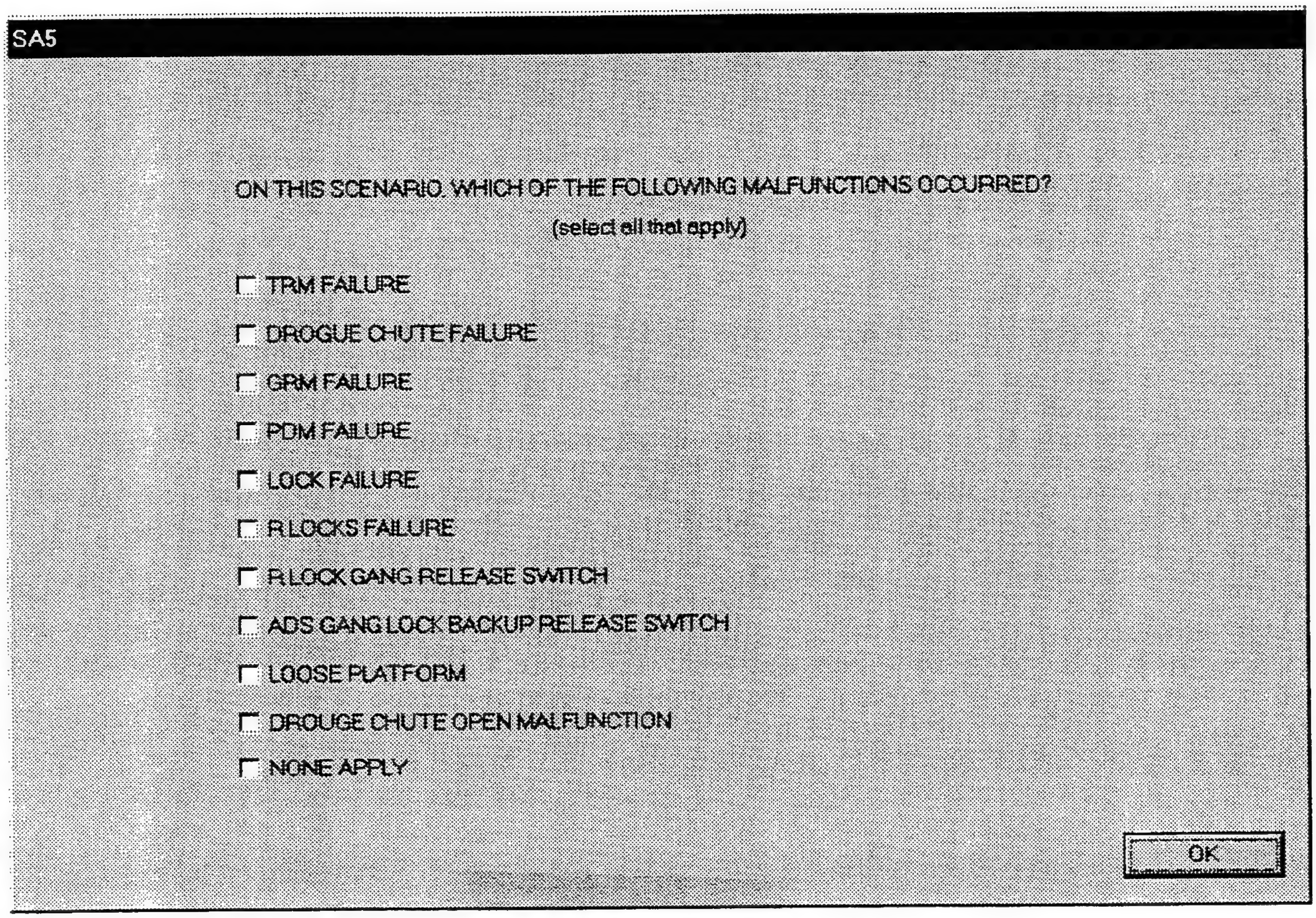
APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



SA5

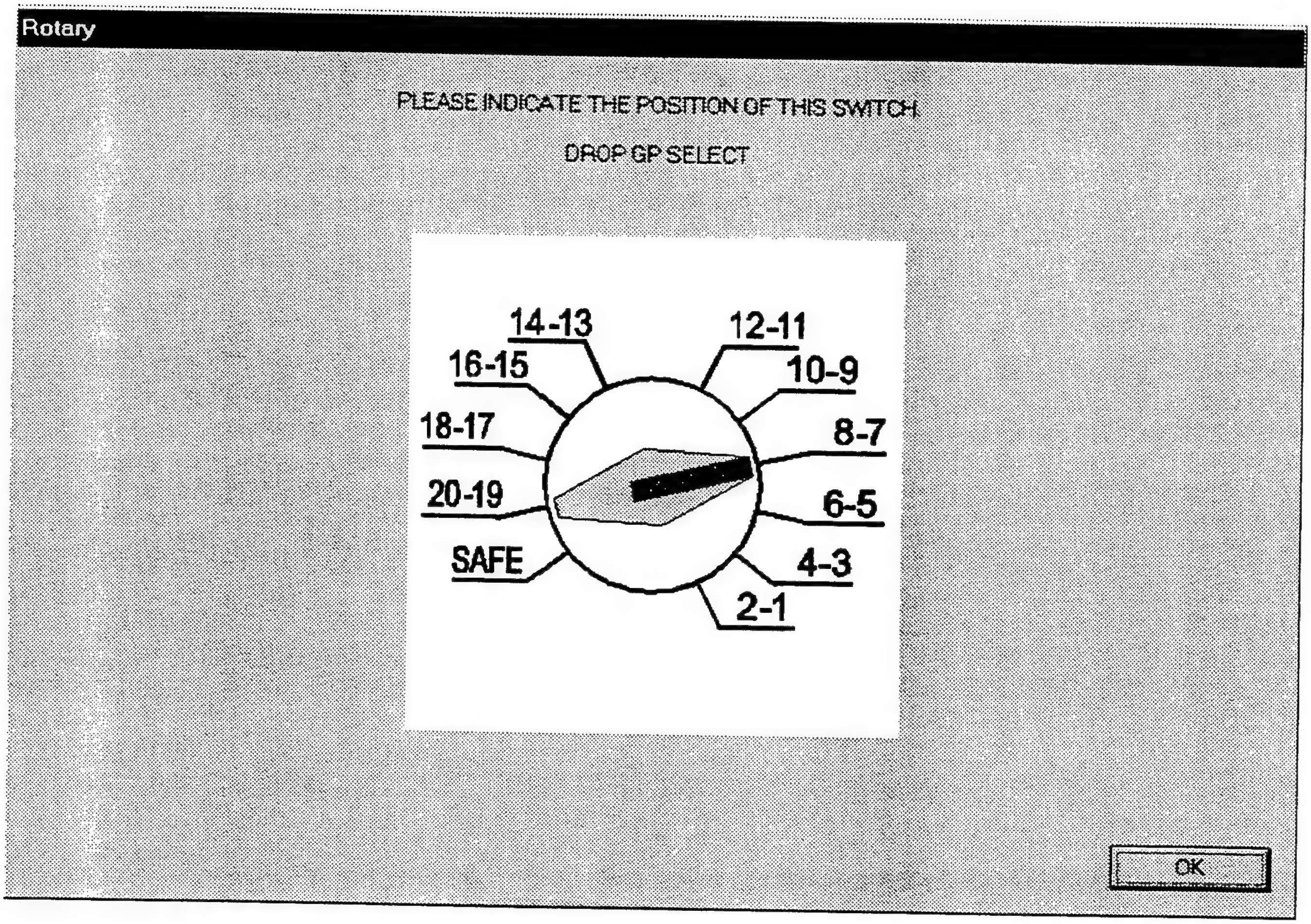
ON THIS SCENARIO, WHICH OF THE FOLLOWING MALFUNCTIONS OCCURRED?

(select all that apply)

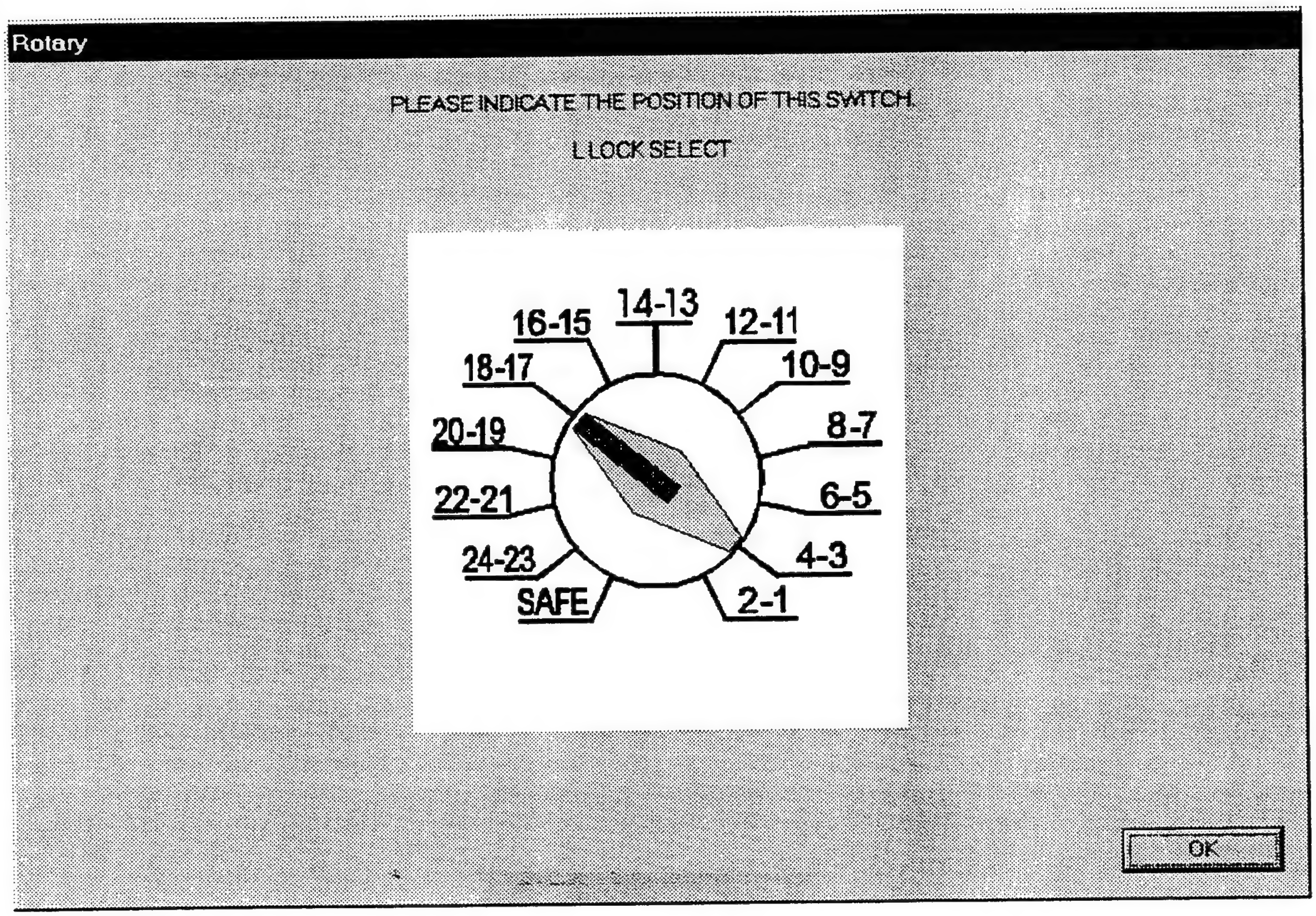
- [ ] TRM FAILURE
- [ ] DROGUE CHUTE FAILURE
- [ ] GRM FAILURE
- [ ] PDM FAILURE
- [ ] LOCK FAILURE
- [ ] R LOCKS FAILURE
- [ ] R LOCK GANG RELEASE SWITCH
- [ ] ADS GANG LOCK BACKUP RELEASE SWITCH
- [ ] LOOSE PLATFORM
- [ ] DROUGE CHUTE OPEN MALFUNCTION
- [ ] NONE APPLY

OK

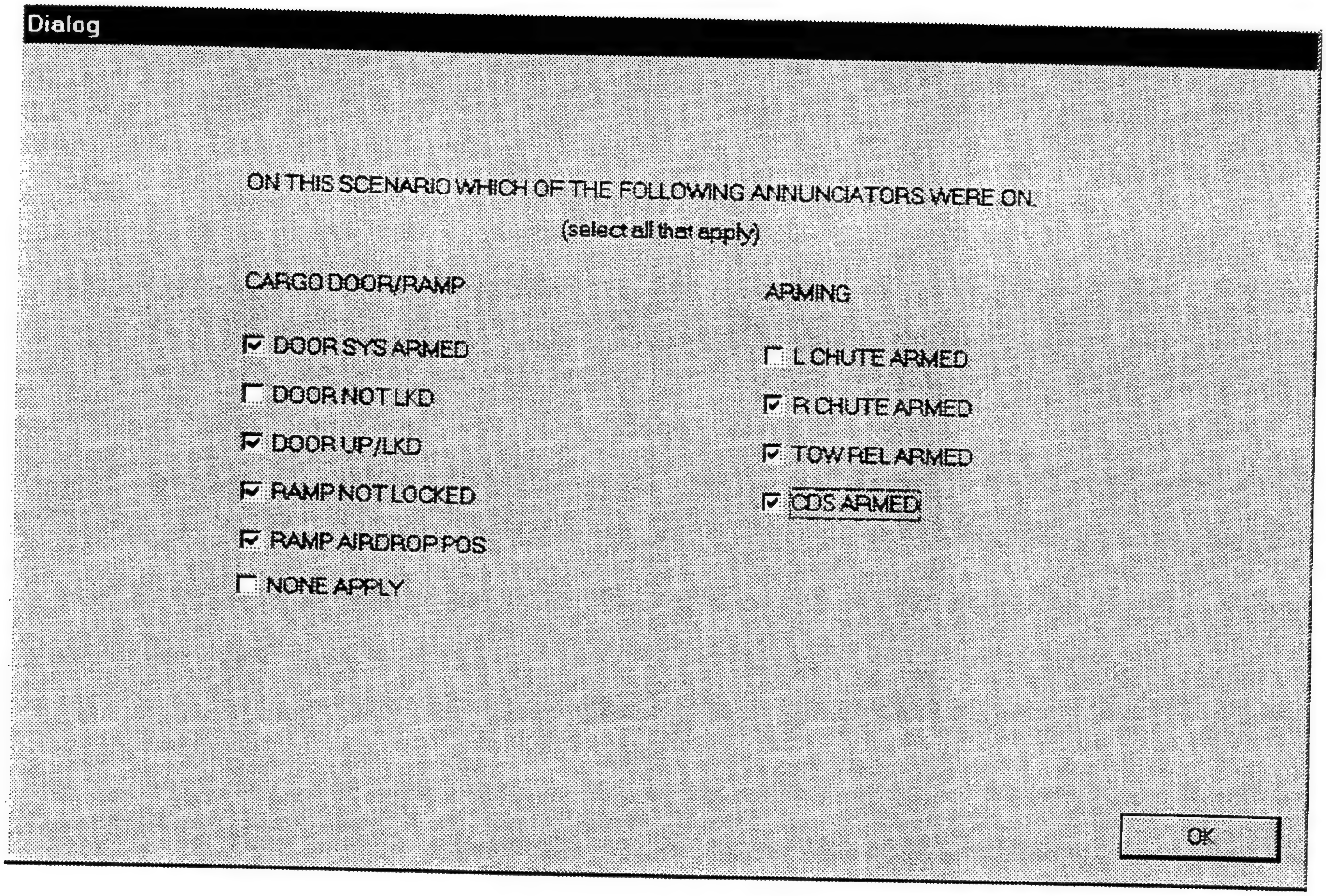
APPENDIX D (CONT'D)



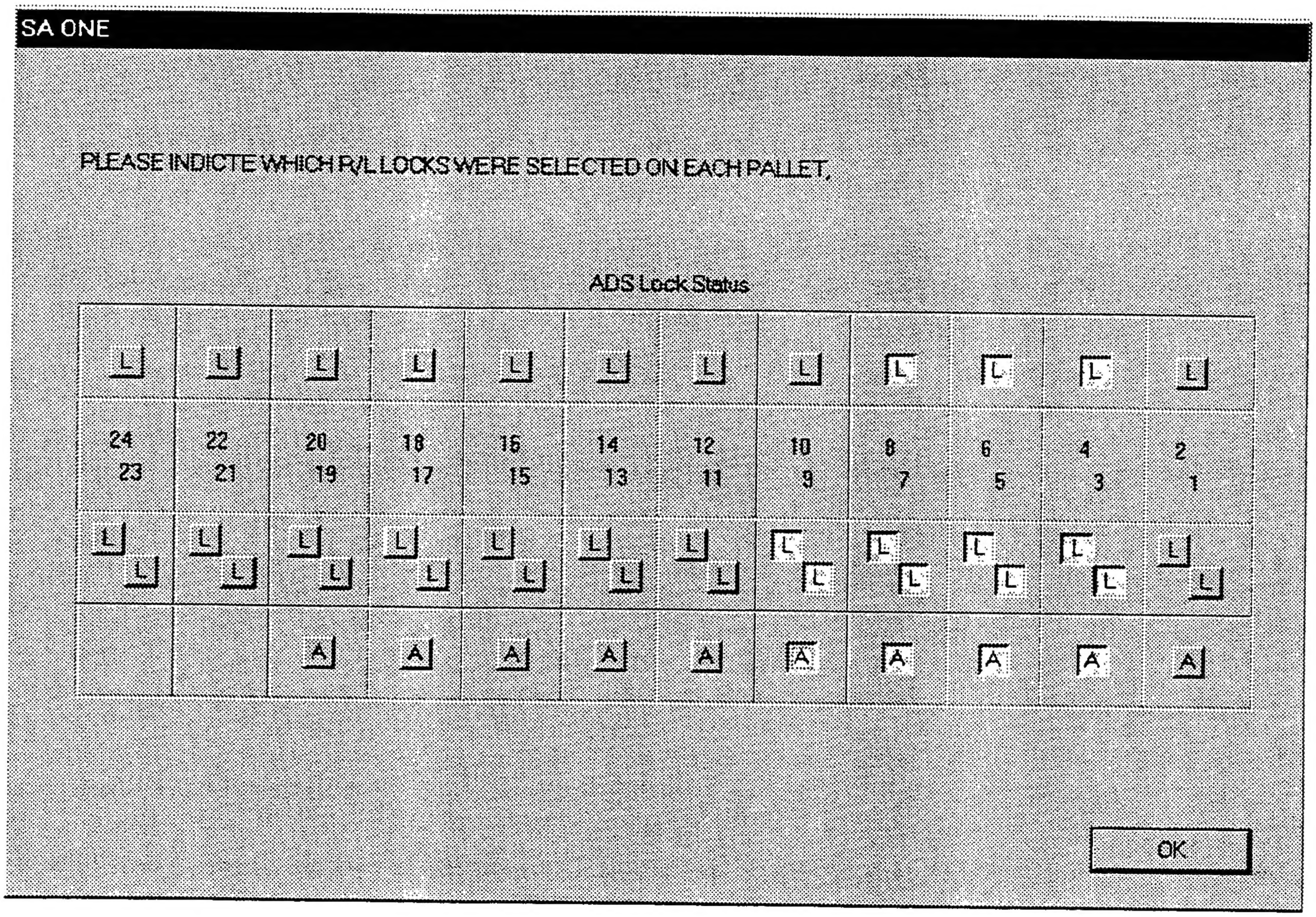
APPENDIX D (CONT'D)



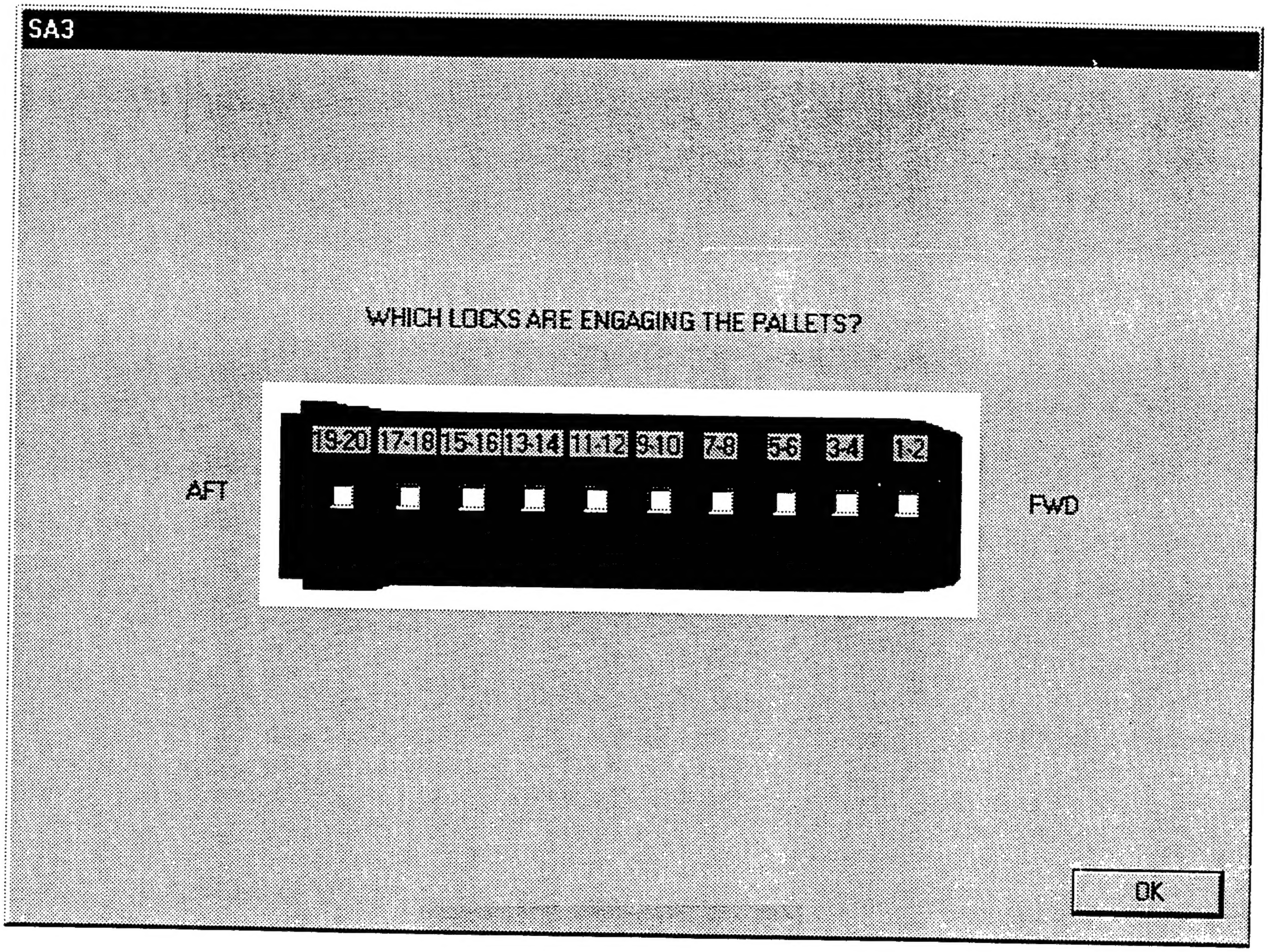
APPENDIX D (CONT'D)



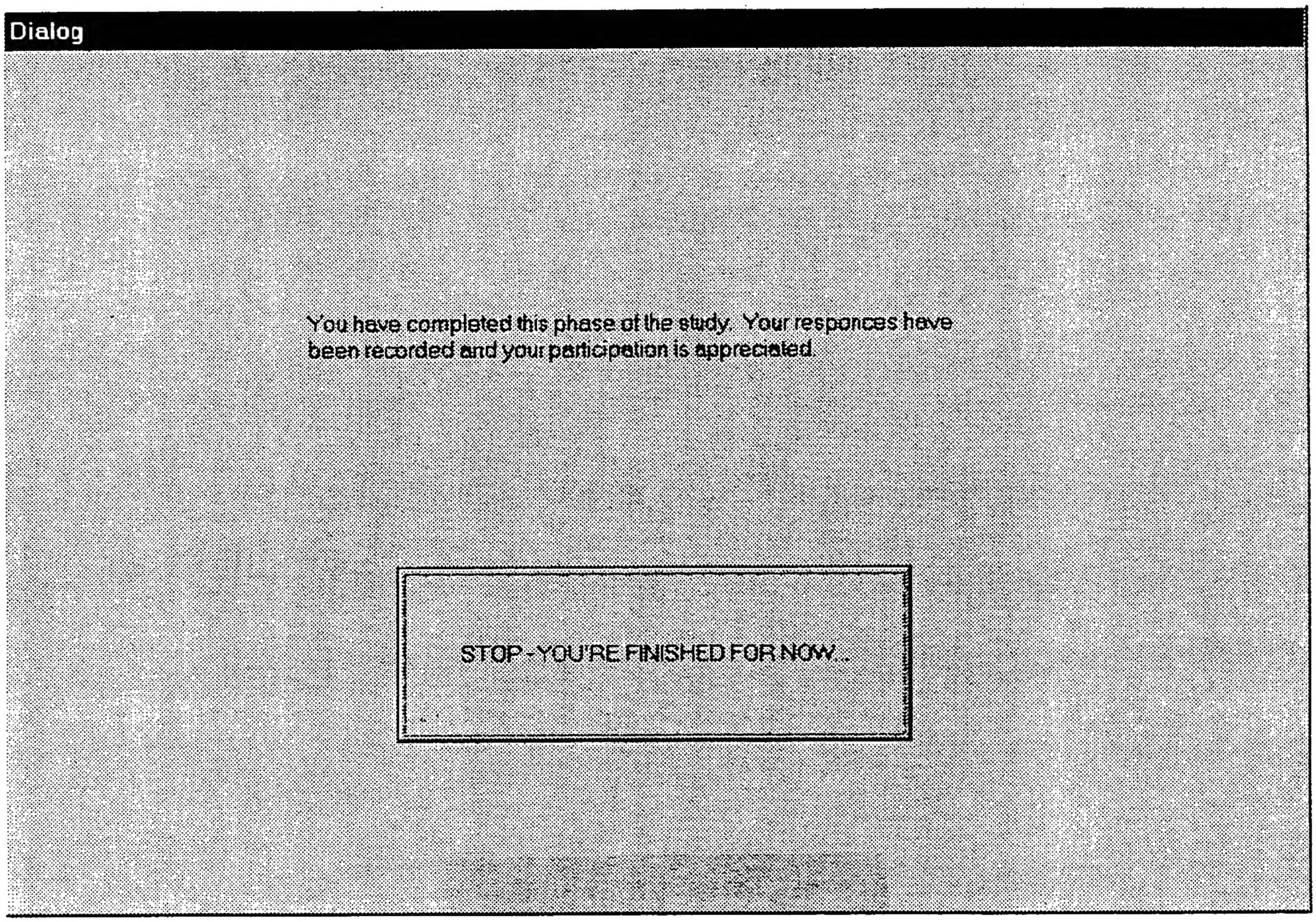
APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



APPENDIX D (CONT'D)



APPENDIX E

SCALE ANCHORS AND DEFINITIONS USED IN SECLECTION OF TEST SCENARIOS

For the scale of involvement, complexity and error, scale points are defined as:

Extremely Low—the task can be accomplished with no or almost no effort;

Quite Low—the task can be accomplished with little to some effort;

Slightly Low—the task can be accomplished with some to a good deal of effort;

Slightly High—the task can be accomplished with a good deal to a great deal of effort;

Quite High—the task can be accomplished with a great deal to tremendous effort;

Extremely High—the task can be accomplished but only with the most tremendous effort.

For the scale of recovery, the reverse would apply.